PSYCHOLOGICAL REVIEW

ological

EDITED

RBERT S. LANGFELD

RGE A., UNIVERSITY OF
TER, CLARK UNIVERSITY
ANGFELD, PRINCETON
. ROBINSON, YALE U
OSEPH PETERSON,

Exper. Psychol.)

(Monographs)

fields and the ng Process

BY

J. A. GENC

UNIVERSITY OF CALIFO GELES

ED FOR THE AMERICAN
SYCHOLOGICAL R
PRINCET
AND ALBA

ATION BY
ANY

of the

ical Associati

SYLVANIA (*J. of Exp. Ps*
TY (*Index and Abstracts*)
(*J. Abn. and Soc. Psycho*
ON UNIVERSITY (*Review*)
UNIVERSITY (*Bulletin*)
Y COLLEGE (*Monographs*)

Business Editor

VIEW
-monthly, January, March
s comprising a volume of

LLETIN
, psychological news and
monthly (10 numbers), the
sues of the BULLETIN co
ment of psychology.

L PSYCHOLOGY
tal character, appears bi-mo
December, the six numbers

INDEX
graphs, and articles upo
during the year. The
bed for in connection

TRACTS
ar supplement making a
o the publication of non-
y and closely related sub

OGRAPHS
tions of laboratory studies
ts. The price of single
s appear at irregular
es.

OCIAL PSYCHOLOGY
March, the four numbers
tains original contributions
s, notes and news.

ON RATES
per volume.
: $6.00 per volume (Foreign,
ts: $6.00 (Foreign, $6.25).
, $5.25). Single copies $1.
1.00; Abstracts, 75c; Bulleti

ATES
, $10.50).
, $11.50).
$12.50).
Foreign, $16.75).
$19.00 (Foreign, $19.75).

ications should be sent to

EW COMPANY

N. J.

VOL. XLV
No. 4

PSYCHOLOGICAL REVIEW PUBLICATIONS

WHOLE NO. 203
1934

Psychological Monographs

Brain Fields and the Learning Process

BY

J. A. GENGERELLI

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

PUBLISHED FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION BY

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.
AND ALBANY, N. Y.

TABLE OF CONTENTS

BRAIN FIELDS AND THE LEARNING PROCESS

BY

J. A. GENGERELLI

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

It requires presumption, if not considerable foolhardiness, to advance at the present time any neurological considerations in an effort to unify some of the many and varied phenomena which constitute the field of learning. In this respect we are suffering from a reaction to the unbridled physiological speculations in which psychologists even slightly interested in this highly important field of study have indulged.

Unfortunately, the impatience of the present day psychologist with any further attempts to "neurologize" is amply justified. Only too often the neurological concepts have been far too naïve,—and necessarily so in the absence of any well-founded knowledge of the physiology of the nervous system. What is perhaps less excusable is the fact that the psychological assumptions were equally naïve.

It should be to the everlasting credit of the members of the Gestalt school of psychologists that they have persistently and insistently pointed out, among other things, certain elementary yet indisputable facts of a dynamic nature which had been either implicitly or explicitly ignored by the more atomistically and mechanically minded psychologists of the late nineteenth century. We refer to that body of phenomena which may be designated by the general class name of "motivation." Köhler, Koffka and Lewin have frequently indicated the flagrant omission of the systematic rôle of motivation from both classical and contemporary psychological theory—a fact which has invalidated the whole frequency-recency hypothesis of association. Since most of the extant neurological speculations were founded upon faulty psychological constructions, it was inevitable that the former

should also be discarded as worthless. Furthermore, the many observed phenomena in the field of learning appear so hopelessly complex when compared to the known facts of physiology of the nervous system that further emphasis has been given to the opinion that neurology is more of a hindrance than a help to systematic psychological progress.

However, within recent years significant advances have been made on both the psychological and neurological sides. In the former domain we have, among others, the outstanding studies of Tolman and his students in animal psychology and the papers from the University of Berlin; in the latter, we have the very ingenious and successful experiments of Adrian, Zottermann, Forbes and Lapicque, to say nothing of many others. In the light of this newly accumulated insight, it might not be amiss to advance some considerations aimed at bringing a certain amount of unity among some of the apparently unrelated facts in the field of learning.

We shall forthwith present the skeleton framework of a possible neurological theory, and begin by listing a number of necessary postulates which we shall employ during the course of the discussion.

1. The afferent excitation entering the cerebral mass diffuses throughout certain regions of that mass, and remains in the form of an after-excitation for some time after the stimulus has been removed.
2. The intensity of the after-excitation varies as some inverse function of the time which has elapsed since its inception.
3. For a given set of cortical conditions, the locus, intensity and spread of excitation and after-excitation in the cerebral mass are unequivocally related to the locus, pattern and frequency of the incoming afferent impulses.
4. The degree of polarization present influences the conductivity of the junction (synapse) between two juxtaposed neurones; and the passage of a train of impulses over a given synapse may alter the degree of polarization.
5. The synapse is a point of increased resistance, and the frequency which it transmits is never greater than the frequency transmitted to it by the centripetal fiber.

I

Fixation and Extinction

We shall consider first in a purely formal manner what, for the lack of a better name, may be termed the processes of *fixation* and *extinction* of response. We may define the former as a tendency to an increase in the intensity of a given response as a result of practice, when a stimulus of constant physical intensity is applied. The latter we may define as a tendency to a decrease in the intensity of the response, under the similar conditions of practice and constancy of stimulus intensity.

The formal problem of determining the differential conditions which make now for the one and now for the other may be safely considered as the crux of learning theory.

We shall begin our discussion by postulating the simplest possible conditions. It would be quite beside the point to object at this juncture that the actual conditions within the very complex organism are not as simple as those we are considering here. We shall inject into our discussion later the various complicating factors contributed by the morphological complexity of the central nervous system as we know it; but for the present we shall have to cling to an *als ob* hypothesis for the sake of simplicity—as is done so often in science.

The consensus of present day opinion among neurophysiologists seems to be that the functioning of nerve may be most adequately conceived under some form of membrane theory of interface potentials. Thus it is commonly assumed that the nerve consists of a protoplasmic core surrounded by a semipermeable membrane and lying in tissue fluid. In the core itself we have a process of electrolytic dissociation which produces ions of opposite signs in equal numbers. If it is assumed that the membrane is permeable to the ions of one sign and impermeable to those of the other sign, it is evident that as a result of the rapid vibratory motion of the ions there will be formed a Helmholtz double layer;

that is, ions of one sign will collect along one side of the membrane while along the other side will be collected ions of the opposite sign. Thus the membrane becomes polarized. A great number of experimental results indicate that the outside of the membrane is positively charged, while the inside is charged negatively.

Authorities such as Lille and Gerard (23, 13) regard stimulation of nerve fiber as being essentially a process of depolarization of this membrane. To quote Troland: "It [the depolarization of the membrane surface] harmonizes . . . with Pflüger's laws of electrical stimulation, according to which there is stimulation at the negative pole (cathode) upon closing the circuit and at the positive pole (anode) upon opening the circuit. Both of these poles are applied to the outside of the nerve. The cathode may be regarded as spraying the nerve with electrons, which are negatively charged and which consequently combine with the positive particles on the outside of the nerve, neutralizing them and thus depolarizing the membrane. The anode, on the other hand, sucks electrons away from the outside of the cell, thus increasing its production of positive particles. Facts to be considered below lead us to suppose that, under these conditions, the membrane will react in such a way as to compensate for the increased polarization, so that, when the circuit is opened and the auxiliary polarizing force of the anode is removed, a depolarizing action will occur. It is a well recognized fact that anode stimulation at the 'break' occurs less readily than cathode stimulation at the 'make'." (31, p. 22.)

The contention that nerve stimulation is primarily a matter of depolarization is further supported by the phenomena of electrotonus. Thus the more recent work of Erlanger, Bishop and Gasser corroborates the earlier work on this subject, showing that at the region of the anode the stimulation threshold is distinctly higher than at the cathode (7, pp. 633–638). It is also indicated by the very careful experiments of the above authors that in the region of the anode the nerve current is of greater amplitude, in the region of the cathode, of smaller amplitude, than the normal action current. This is what we should expect

if we regard stimulation of nerve fibers as being a matter of the reduction of polarization beyond a critical limit; for evidently the positive charge at the cathode will be less, hence the potential difference smaller; at the anode the positive charge will be greater, and the potential difference correspondingly larger.

The mechanism of propagation does not concern us with regard to our immediate problem and we may therefore forego any remarks upon this topic except to indicate the findings of Erlanger *et al.* (*loc. cit.*, p. 639) that as the amplitude of the nerve current decreases under the cathode, the rate of conduction increases; as the amplitude increases when a propagated wave reaches the anode, the rate of conduction decreases. In other words, the smaller the amount of positive charge (within limits) on the outer surface of the membrane, the greater the speed of conduction; the greater the amount, the less the conduction.

Let us now suppose, in accordance with our deliberate policy of artificial simplicity, that we have two neurones, x and y, which are in synaptic juxtaposition, and that the direction of current flow is from x to y. Turning our attention to the conditions at the synapse, we at once perceive the probability that this region is a locus of increased positive charge. This accentuated positivity is obviously brought about by the juxtaposition of the two positively charged surfaces, the end brush of the neurone x and the dendrites of neurone y. This increased positivity may be properly considered as common to the two neurones at that point.

Now we have seen that increased positive charges at the outer surface of the fiber membrane serve to increase the stimulation threshold and diminish rate of conductivity. Consequently we may justifiably expect the synapse to present a region of increased "resistance" and diminished conductivity. It is, in other words, a region presenting a state of affairs analogous to that prevailing at the anode (**31**, p. 49). A single impulse arriving at the synaptic region from the axon of x will, therefore, by itself be probably incapable of sufficiently depolarizing the synaptic interface to enable the "propagated tendency to excite" to continue and effect a stimulation of the adjacent neurone. If, however, several impulses, coming sufficiently close together, invade the

synaptic region, it is clear that, as a result of summation, the interface will be critically depolarized and an impulse will be instituted in the dendrites of neurone y (yd). It is thus evident that the rhythm of discharge instituted in yd will not correspond to the rhythm of discharge coming to the synaptic region from the axon of neurone x (xa).

Before proceeding to a consideration of the possible quantitative relationships involved in the above process, it might not be amiss to say a few words anent the "polar" nature of conduction at the synapse. The following considerations occurred to us as being pertinent to a possible explanation of the irreversible quality of conduction at synaptic junctions.

It is a familiar fact in neurophysiology that the irritability of nerve fibers varies as some direct function of their diameter. Since the end arborizations of axons are in general smooth and thin while dendrites are characteristically thick and spiny, we may have in this structural differentiation the basis of an explanation for synaptic polarity. On these grounds, it is therefore conceivable that an end brush may stimulate the more sensitive dendrite; but not *vice versa*. We were gratified to discover that this hypothesis had been advanced by Forbes (8). We quote Gerard: ". . . If a small and large diameter fiber were in contact, it is conceivable that an impulse traveling along the small fiber might successfully pass on into the large, which has a lower threshold, whereas the reverse might not be true. The junction of a fine and coarse fiber, then, would account very simply for the irreversible conduction at the synapse (13, p. 75).

It is a misfortune for psychological theory that so very little is known about the biochemistry of nerve function. Any quantification which may be envisaged to account for the diverse phenomena of modifiability becomes, as a consequence, nothing more than rule of thumb hackwork. Any mathematical equation which may be developed to describe processes at the synapse can be judged, therefore, not on the basis of its accuracy in reflecting the microscopic events which go on there while the nerve is functioning (since these are unknown), but on the somewhat negative

grounds of its adequacy in embracing a wide variety of facts on the psychological side, without doing injustice to such few well-ascertained neurophysiological facts as we do possess. Hence it must lack that deductive elegance which is so dear to the hearts of theorists.

Returning to our two juxtaposed neurones x and y, let us assume that the axone of neurone x (xa) is contributing to the synapse a given sufficient frequency f_{xa}, and that the two neurones have never conducted heretofore. Then an impulse frequency will be set up in the dendrites of neurone y (yd). This latter frequency f_{yd} or f_s we believe to be less than f_{xa}. (See postulate 5, and discussion in Section 2.) Now, if we assume that this original volley f_{xa} has brought about a given amount of diminution in polarization at the synapse which endures for a period of time, the next volley of impulses over xa of the same frequency value will conceivably institute a change in the frequency with which yd will respond. It becomes our duty, therefore, to determine and evaluate the factors which make for change in the amount of polarization at the synapse. Obviously the amount of positive charge at the synapse may vary in two directions. It may either increase or diminish. If it increases, we have as a consequence of our preceding considerations a diminution in the irritability or conductivity of the synapse; if it diminishes, we have an increase in irritability and conductivity. Clearly, the one process moves toward extinction; the other, toward fixation. Further, if the process moves toward increase of polarization at the synapse, we should expect the value of f_s to diminish with each successive bombardment of the synapse by the f_{xa} volley (we assume f_{xa} to remain constant throughout); whereas, if the process moves toward a diminution of polarization at the synapse, we should expect the value of f_s to increase toward some limit with each successive bombardment.

The critical reader may well object here that there is not a shred of experimental evidence to justify the assumption of such a relationship as is expressed in the preceding paragraph. He might very well argue that since nerve impulses are the only agents present which are in a position to increase or decrease the

amount of positive charge at the synapse, these same impulses should have the same or, at all events, a similar, effect on the external membrane surface of the axon. That, certainly, an increase in nerve impulse frequency in axons to faradic stimuli as a result of a previous application has never been observed; and that since, on the basis of our contention, the synapse differs from the axon itself primarily in degree of polarization, there is no reason to believe that such a relationship should hold true at the synapse any more than at the axon.

There are two points to be made in reply to this objection: First, the effect in question, if it exists in the axon, would not show itself except under certain special conditions, and to our knowledge, these conditions have never been met in experiment. It is a well known fact that faradic stimulation of a nerve has an upper threshold frequency beyond which the nerve ceases to respond. If now, we stimulated such a nerve briefly and at *very* short intervals with a faradic frequency near its upper threshold, we should, if the above-mentioned effect appears, obtain an increase in the upper threshold frequency. (See *infra.*) In other words, we should obtain a diminution of the refractory period. Such an experiment, however, would be beset with difficulties, owing to the possible distorting effects of the myelin sheath.

That this effect may be well within the realms of possibility is indicated by the results of Erlanger, Bishop and Gasser in their work on the polarization of a nerve. We quote: "For weak cathodal polarization, there occurred a decrease of amplitude of one-fourth to one-third, an insignificant increase of duration of rising phase of .33 σ to .34 σ, and a decrease of duration of the falling phase of 1.29 σ to 1.06 σ, or 82 per cent of normal." (*Loc. cit.*, p. 647.) In other words, since cathodal polarization presumably diminishes the amount of positive charge at the external membrane surface, we may believe that this effect serves to diminish the total temporal value of the falling phase of the bioelectric variation. And although there is no unanimity of opinion among nerve physiologists as to the relationship between the duration of the falling phase of the impulse and the size of

the refractory period, there is some evidence that some form of direct relationship exists.

In the second place, while we consider the synaptic interface merely as a region of increased polarization, it is almost certain that some effects should take place there which might be less marked or absent in the axon. For, in addition to the fact that nerve fiber is unmyelinated at the synapse, we have the probable fact of the changing spatial relationships of dendrites and end-brush to each other at that point as a result of conduction. Thus Kappers' Law of Neurobiotaxis, that juxtaposed axons and dendrites grow more closely together provided that these two parts of neurones are in excitation simultaneously or in close succession may be the direct outcome of the polarization process at the synapse consequent upon a volley of impulses passing through it. For, obviously, if the two surfaces at the synapse bear charges of the same sign, they will mutually repel and thus a tension is created which tends to separate the two endings from each other. During conduction through the synaptic interface this repulsion is somewhat dissipated, and the dendrites and end-brush are freer to effect a better *rapprochement*. Furthermore, as a result of this *rapprochement,* we shall see that repolarization cannot reach its original intensity because of the greater physical contact between the two units; hence the *durational* aspect of the modification may be understood.

Let us consider, therefore, what possibly might happen at the synapse when impulses pass over it. Since the nerve membrane is positively charged during a resting condition and the nerve impulse seems to be a propagated wave of negative potential accompanying a similar wave of temporarily increased permeability in the membrane, it follows that the negative ions in the nerve core will surge forward at right angles to the surface of the membrane. As they pass through, some of them will be captured and held by the membrane itself, others will unite with and neutralize the free positive particles present at the outer surface of the membrane. At the same time, however, it must be remembered that the negative ions are not the only ones which leave the core and penetrate through the membrane while it is

in a state of temporary increased permeability. There will be positive ions as well, although probably not in such great numbers, for there is no reason to believe that all the positive ions formed in the core during the period of electrolytic dissociation will find their way to the external surface of the membrane. We must consider, then, that at the synapse with each impulse there are liberated toward the outer surface of the membrane both negative and positive ions, and the net depolarization effect there will be the algebraic resultant of these two quantities.

It is highly probable, moreover, that during the first few impulses relatively few of the negative charges thus liberated will succeed in taking their place along with the free positive ions which constitute the external surface layer. For, during the resting state of the nerve, some of the accumulated free positive charges will have been adsorbed by the membrane itself and consequently it will have become positively charged. Hence, most of the anions will be captured by the ionized molecules of the membrane. As more and more impulses pass over the nerve fiber, however, the number of negative ions thus captured by the membrane itself becomes smaller and smaller. This leaves an increasing number of anions free to neutralize the free positive charges.

It is interesting to note that the above purely theoretical considerations have conceivable support in the phenomenon of *after-potential,* studied by Gasser and Graham. We quote from their recent article: " During the last 85 to 90 per cent of the visible duration of the spike potential, the nerve shows signs of the production of another potential very different in its general behavior, the *after-potential.* The after-potential has a rising phase. Like the total duration, this rising phase lasts for very variable times, according to the condition of the nerve. Its crest may occur at 3.0 to 4.0 σ at the lower limit up to 60 σ or more at the upper " (9, p. 330). The duration values of the after-potential are probably distorted by the effects of the myelin sheath. But it would be interesting to determine experimentally whether the extent and sign of the after-potential varies with the

frequency of the impulses, as would be expected from our theoretical considerations. (See *infra.*)

The considerations in the preceding paragraph but one tacitly assume that there is no decay or dissipation of the negative charges captured by the membrane. This probably is not the case. Hence we must consider that the dissipation process will tend to undo what the absorption process, through the propagation of the impulse, has accomplished. If we assume with Blair (2, p. 726) that this process of decay is exponential with time, it becomes at once evident that the depolarizing process at the synapse will depend upon the frequency of the nerve impulses (f_s) which are conducted through it. For, if the nerve impulse frequency is sufficiently slow, the effect of the absorption by the nerve membrane will be neutralized by the "decay" and the succeeding impulse will again fail to contribute to the surface of the membrane its share of negative ions.

Let us set

θ, the amount of positive charge present at the synaptic interface.

θ_1, the amount of positive charge released by a single impulse.

ϕ_1, the amount of negative charge released by a single impulse.

f_{xa}, impulse frequency with which x neurone bombards synapse.

t, duration of impulse volley coming over neurone x.

L, limit of amount of negative charge which membrane will capture and hold.

$L=\phi_1$

$\phi_1>\theta_1$

By Blair's law of decay, the amount of negative charge retained by the membrane when the next impulse arrives is

$$\frac{\phi_1}{e^{\frac{b}{f_s}}} \quad \text{or} \quad \frac{\phi_1}{e^{bt_1}}$$

if we set $1/f_s=t_1$. b is a constant. The amount of negative charge which this impulse will release to the synaptic interface will be

$$\phi_1-(L-\frac{\phi_1}{e^{bt_1}}) \qquad (1)$$

which, since $L=\phi_1$, becomes

$$\frac{\phi_1}{e^{bt_1}} \qquad (2)$$

But with each impulse, positive ions to the extent of θ_1 are also released toward the membrane surface, and these will have a corresponding neutralizing effect. Therefore, the actual amount of negative charge free to affect the synaptic interface is given by

$$\frac{\phi_1}{e^{bt_1}} - \theta_1 \qquad (3)$$

It is evident, therefore, that the effect of activating a synapse may result in *either an increase or a decrease in the amount of total positive charge there.*

We may now attempt to follow the processes at the synapse as they are affected by the frequency of the impulses passing over it and by the duration of the volley. We assume that the axone of neurone x is conducting to the synapse with a constant frequency f_{xa} which is sufficient to arouse some frequency f_s in the synapse, transmitted to neurone y. It is simpler and more rational to conduct our discussion in terms of the temporal intervals separating the successive impulses crossing the synapse; that is to say, in terms of t_1. For sake of simplicity, we will assume that the duration of the bioelectric variation of a single impulse is constant and negligible.* Now let us write

$$t_1 = \frac{a\theta}{F_a(f_{xa})} \qquad (4)$$

where a is a constant, and F_a is some direct function.

* This assumption, while obviously unjustified, does not alter the tenor of our considerations and conclusions. It is highly probable that duration of the bioelectric variation is influenced by the amount of positive charge at the synapse itself. *E.g.*, Erlanger *et al.* report: "For weak anodal polarization there was . . . an insignificant increase in the rising phase of the impulse of $.33\sigma$ to $.43\sigma$ or $.10\sigma$; and an increase of duration of the falling or restoration phase of 1.29σ to 2.22σ" (*loc. cit.*, p. 647). It is also highly probable that the values of ϕ_1 and θ_1 are a function of the duration of the bioelectric variation itself, and therefore are not constant, as we have assumed. But rather than involve ourselves in these complications, we prefer an artificial simplification to give us such insight as may be obtained from such a premature inquiry.

According to this expression, the temporal interval separating any two successive impulses initiated in the synapse at any given time by the conducting centripetal fiber varies directly as the amount of positive charge present at the synapse at the beginning of the interval t_i and inversely as some function of the impulse frequency with which the centripetal fiber is conducting. We may look upon the constant a as a value which is peculiar to a given synapse and expresses the influence of the chronaxie of the nerve tissue involved and any other conditions which are present and relevant.

It is at once evident, however, that this expression is not complete. We have neglected the influence of adaptation in nerve tissue. Judging from what is known of the adaptation effect in nerve, we may expect that this factor tends to make for diminished irritability; that is to say, it will tend to increase the value of t_i. Therefore we should write

$$t_i=\frac{a\theta}{F_a(f_{xa})}+A \qquad (5)$$

where A is the amount of adaptation incurred for a given duration of the volley.

We must now attempt to evaluate A. If we take our cue from Hecht's analysis (15) with photoreceptor processes, we may propose some differential equation as the following:

$$\frac{dA}{dt}=k(cf_{xa}-A) \qquad (6)$$

where k and c are constants, and t is the time which has elapsed since the beginning of the impulse volley. Integrating, we obtain

$$A=cf_{xa}-e^{-kt} \qquad (7)$$

Equation (5) now becomes

$$t_i=\frac{a\theta}{F_a(f_{xa})}+(cf_{xa}-e^{-kt}) \qquad (8)$$

We are now in a position to follow the depolarization (or polarization) process at the synapse.

We will start with an amount of positive charge at the synapse equal to θ_o. With the axon of neurone x conducting at a fre-

quency f_{xa}, a period of time t_o may be considered to elapse before the synapse responds with its first impulse; and a period of time t_1 may be considered to elapse before the synapse responds with the second impulse.

From our considerations in the preceding paragraphs, it is evident that the result of the first synaptic impulse will be to increase the extent of positive charge at the synapse by the amount θ_1. The amount of positive charge at the synapse at the moment when the second synaptic impulse is initiated will then be (neglecting any effects of diffusion) $\theta_o+\theta_1$. It now remains to determine what will be the effect of the second impulse on the polarization state of the synapse. This may be determined by substituting in equation (1), and as a result we have

$$\frac{\phi_1}{e^{bt_{11}}} \qquad (9)$$

Where t_{11} is the interval separating the first and second synaptic impulses. If the interval t_{11} is sufficiently small, we shall have a balance of negative ions released to the surface, approaching $\phi_1-\theta_1$ (approximately) as a limit. If t_{11} is sufficiently large, we shall have a balance of positive ions released toward the membrane, approaching θ_1 (approximately) as a limit. It is clear, then, that whether the synaptic processes will move toward increased or diminished polarization and therefore increased or diminished "resistance to synaptic conduction" is determined by the size of this first interval t_{11}.

Let us assume that, in this case, the value of t_{11} is such that there will be a balance of negative ions, and set this amount equal to ϕ_{11}. Then, obviously, the amount of positive charge at the synapse will be diminished as a result of the second impulse, and will be

$$\theta_o+\theta_1-\phi_{11}$$

(Had there been a balance of positive charges, the value $\theta_o+\theta_1$ would have been *increased* by an amount determined from the relationships expressed in equation [1].)

We are now ready to determine the value of t_{12}, or the interval which will separate the second and third impulses. This will be

$$t_{i2}=\frac{a(\theta_o+\theta_i-\phi_{i1})}{F_a(f_{xa})}+(cf_{xa}-e^{-k(t_o+t_{i1})}) \quad (10)$$

Having obtained t_{i2}, we repeat the above process and determine the effect on synaptic polarization which the third impulse contributes. If this is again in the nature of a balance of negative ions, the value of t_{i3} will be determined by

$$\frac{a(\theta_o+\theta_i-\phi_{i1}-\phi_{i2})}{F_a(f_{xa})}+(cf_{xa}-e^{-k(t_o+t_{i1}+t_{i2})}) \quad (11)$$

If, on the other hand, the result of the third impulse is the contribution of positive ions to the synaptic interface, then

$$t_{i3}=\frac{a(\theta_o+\theta_i-\phi_{i1}+\theta_{i2})}{F_a(f_{xa})}+(cf_{xa}-e^{-k(t_o+t_{i1}+t_{i2})}) \quad (12)$$

And so on, until we come to the end of the duration of the impulse volley.

If we substitute appropriate constants in our equations and give F_a a definite value, we may follow the processes in the laborious fashion which has been indicated. By so doing, it may be shown that the synaptic conditions may follow any one of the three following courses (depending upon the values of f_{xa}, θ_o, ϕ_i and θ_i), throughout the duration of the impulse volley.

1. The amount of positive charge at the synapse will decrease, first at a positively accelerated rate, then at a negatively accelerated rate until a limit is reached; whereupon it will increase, first at a positively accelerated rate for a short time, and then at a negatively accelerated rate as it approaches a limit.

2. The amount of positive charge at the synapse will decrease at a negatively accelerated rate until a limit is reached; whereupon it will increase, first at a positively accelerated rate for a short time, and then at a negatively accelerated rate as it approaches a limit.

3. The amount of positive charge will increase at a positively accelerated rate for a short time, and then at a negatively accelerated rate as it approaches a limit.

The three types of process may be illustrated by Figure 1, where curves 1, 2, and 3 represent the alternatives 1, 2, and 3, respectively.

We may lump together, as it were, curves 1 and 2 and consider them as representing the course of conditions at synapses underlying the phenomenon of fixation; curve 3 we may regard as illustrating the course of synaptic conditions underlying the phenomenon of extinction. The former two curves illustrate

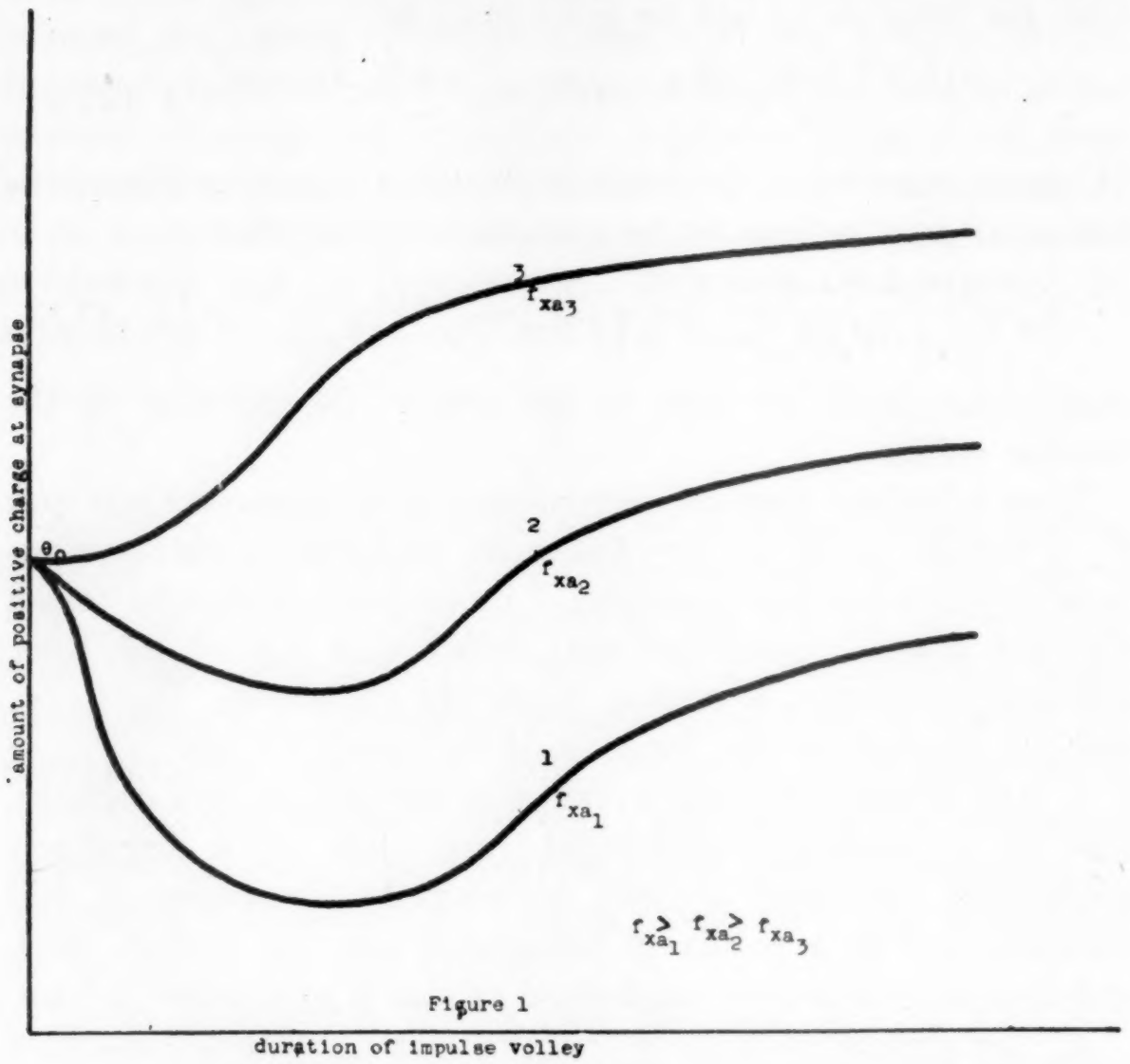

Figure 1

FIG. 1. Variation of θ at the synapse with varying values of centripetal impulse frequency and duration of impulse volley. All other factors are assumed constant.

what happens when the value of f_{xa} is sufficiently great; the latter curve represents the state of affairs obtaining when f_{xa} is too small.

Hence, from the point of view of our considerations, whether the process of modification of reaction will move toward fixation or extinction is determined by the value of the impulse frequency bombarding the synapse (together with the value of θ_o). More

accurately, it is determined by the values which f_{xa} and θ_o assume with relation to each other. It is further evident that even though the value of f_{xa} should be sufficiently high to insure at the beginning a process of de-polarization at the synapse, the process is reversed when the duration of the impulse volley exceeds certain critical limits. In other words, the curves indicate that an optimum duration of stimulation application exists.*

It is clear that thus far we have concerned ourselves with the course of changes occurring at the synapse for the duration of the impulse volley; or, in other words, with the effect of the application to the receptor of a stimulus with a given energy value for finite periods of time.† But since the learning process is consequent upon the repetition of certain stimuli or situations, we must attempt to determine what may happen at the synapse on the second, third and subsequent applications of the stimulus energy. We will assume that stimulus energy is constant, and that the value of f_{xa} and the duration of the volley are likewise constant. (See footnote.)

If we consider expression (8) and assume conditions leading to fixation, it is clear that with each successive application of the stimulus, the amount of reduction in θ becomes greater and greater, provided the duration of the impulse volley is not too long. That is to say, the process of *de*-polarization at the synapse will be positively accelerated. It is highly probable, however, that this state of affairs will not continue indefinitely, and for the following reasons:

As the amount of positive charge at the synapse becomes

* It will be obvious to the most casual reader that we have omitted from our discussion a number of considerations which are most certainly pertinent. Such as: the duration of the bioelectric variation, the influence of diffusion, the effect of adsorption on the behavior of surface molecules in the membrane, the influence of the absorption of negative charges by the membrane on its threshold of "breakdown," the *modus operandi* of the adaptation process, the process of repolarization as it occurs between successive impulses, etc. We have taken merely the simplest possible assumptions and the most "mechanical" relationships and attempted to construct by means of them the barest framework upon which to hang some of our disjointed psychological phenomena. Our ignorance is too great for more.

† We have deliberately ignored the influence of adaptation at the receptor and the receptor-afferent fiber junction, although it is evident that this would serve to diminish the value of f_{xa} which we considered constant.

smaller and smaller, the distance separating the endbrush of the axon of neurone x and the dendrites of neurone y becomes less. (See the considerations on p. 9.) This greater proximity of the two terminations automatically hinders to a certain degree the restoration of polarization at the synapse during the interval separating successive impulses traversing the synapse. The mechanism underlying this phenomenon is to be found in the inverse square formula describing the mutual influence of electrostatic charges upon each other. Clearly, the closer the surfaces of the two neurones come together, the greater the repulsion will be between the positive ions striving to come to the surface of the two membranes, and hence fewer and fewer cations will succeed in invading the synaptic interface in unit time and thus reporalizing it within either the relatively small temporal interval separating two successive impulses or two volleys of impulses when the stimulus is applied at relatively short intervals.*

Since repolarization goes on at the synapse at such an abnormally slow rate, it is manifest that at some point in the series of stimulus applications, the *de*-polarization may reach and surpass a critical value which will serve to inhibit to some degree the arousal of further impulses in the synapse. This may take place in the form of a lengthening of the absolute refractory phase of the impulse owing to the paucity of positive charges in the field. It will be remembered that this is the probable reason why a continuous flow of current applied to nerve fiber results in but one or two impulses; for, obviously, in view of the nature of the nerve impulse, there can be no " breakdown " of the nerve membrane with its concomitant wave of negative potential unless there has been a restoration of the membrane to its original equilibrium; and there can be no restoration of the membrane to equilibrium in the face of an amount of negative charge surpassing the critical threshold. Our suggestion in regard to the influence of synaptic *de*-polarization on the value of the absolute refractory phase has, furthermore, some experimental

* This we consider to be, in the last analysis, the reason that the synapse is the seat of *retained modification*. In the axon, obviously, this process cannot occur owing to the absence of repelling forces, and hence repolarization takes place very quickly.

support. Erlanger *et al.* report that cathodal polarization serves to definitely increase the length of the absolute refractory phase (*loc. cit.*, p. 650).

If we are willing to believe, then, that synaptic *de*-polarization beyond a certain critical point serves to increase the absolute refractory phase by varying amounts, it is clear that once this point is reached and surpassed, the frequency of f_s will become less, and the value of t_i will become greater. If we assume now that the length of the absolute refractory phase is some direct function of the amount of *de*-polarization beyond the critical point, we may well believe that the successive values of t_i will become larger and larger. This, in turn, causes ϕ_i to become smaller and smaller, and hence the total *de*-polarization effect of each successive impulse volley to become smaller and smaller as it approaches the limit zero.

It is clear from the above discussion that the course of *de*-polarization at the synapse in consequence of a series of stimulus applications will be sigmoid in nature. That is to say, the curve describing the fixation process will be sigmoid. This is true, of course, only when the synapse involved in the fixation process is in a sufficiently *pristine* condition. If it has already been activated by other stimuli and depolarized to a sufficient amount, the course will be negatively accelerated. The relevance of this deduction to the current discussions anent the nature of the learning curve is immediately evident.

It is most interesting to note that Kleitman and Crisler in their study of the conditioned salivary reflex, found the process of conditioning to be sigmoid (20).

There is implicit in our considerations another deduction: viz., that, within limits, the wider the temporal spacing between successive applications of the stimulus, the greater the amount of *de*-polarization or fixation. We may persuade ourselves of this by considering that the synaptic adaptation wrought at the synapse by the duration of the impulse volley will require time to become dissipated. Although we have no experimental evidence in this specific connection, it is not unreasonable to assume that the process of synaptic adaptation recovery will be negatively

accelerated. Hence, after the termination of the impulse volley, the irritability of the synapse will increase at a negatively accelerated rate.

This course is complicated, however, by the repolarization process which has the reverse effect. Hence, the actual change in irritability at the synapse during the interval separating successive volleys of impulses will be an algebraic sum of these two factors. Evidently, beyond a certain point, adaptation recovery will be more than offset by repolarization, and at this point, " disuse " begins to exact its toll.

We shall have occasion to utilize this conclusion in a later section.

We have been discussing the course of events occurring in a series which involves fixation. When we assume conditions at the synapse which make for extinction, that is to say, a value of f_{xa} and θ_o which makes for increased polarization from the very beginning, the process will be evidently negatively accelerated, without reversal point.

II

Diffusion Processes in the Cerebral Mass

We have been considering in the previous section a situation which, for the sake of exposition, we have given an artificial simplicity. Obviously, the conditions described above cannot be duplicated in reality. To conceive of nerve currents traversing isolated paths and their synapses has been only too frequently practiced. If deeper insight is to be obtained in regard to the mechanics of neural events, we must avail ourselves of the contemporaneous developments in nerve physiology.

The recent investigations in the field of neurology lead us to believe that the brain is a marvelously intricate conduction network of neurones in which practically every segment of the network is directly or indirectly connected with every other segment. Thus, when a particular receptor is stimulated, the impulse may not travel over an isolated chain of conductors, to eventuate into an efferent fiber and finally into an effector. Rather, by virtue of the network-like nature of the brain and of the cortex in particular, the excitation process diffuses throughout a greater or lesser portion of the cerebral mass. That this is the case has been attested again by the recent work of Travis (30) on one of the more elementary reflexes. The excitation process may or may not eventuate into a response, but the process of diffusion, as it may be termed, exists in all cases. Whether the intensity of this diffusion process varies with the distance from the cerebral locus of excitation, the point of entrance into the cortex, let us say, of the centripetal fibers, it is at present difficult to say from experimental evidence. The problem is one of great theoretical importance, however, and although experiments are lacking—experiments, by the way, which might be readily performed—we are forced to consider it in the light of certain plausible neurological conceptions.

A pertinent fact to be considered in this connection is the

phenomenon of "after-excitation." That is to say, after the stimulus has ceased to excite the receptor, the cortex remains in a state of heightened electrical potential for varying periods of time. To find a reason for the persistence of this heightened potential is of pivotal importance for our subsequent discussion. We need not enter into the discussion of Guthrie (14) and Pavlov (25) on this subject. Guthrie has maintained that this effect is to be looked upon as a form of circular reflex phenomenon, while Pavlov advocates the "central" point of view. As a matter of fact, they are probably both correct in their contentions. Certainly it is a fact that kinaesthetic stimuli consequent upon a reaction contribute to maintain a heightened electrical potential in the cerebral mass when the original stimulus no longer exists; but it may be shown that the "after-excitation" may also exist independently of kinaesthetic reënforcement.

To facilitate consideration of a possible central mechanism underlying this phenomenon, we present the highly idealized diagram shown in Fig. 2. We have pictured here a network of neurones and synapses which, for the sake of simplicity of exposition we have made purely schematic. The arrows indicate the direction of the nerve impulses. Clearly, the drawing expresses the possible relationships between the neurones in two dimensions only. The addition of the third dimension is avoided for the sake of clearness, since the inclusion would add nothing new in principle. Now, if we conceive the neurone relationships in the upper layers of the cortex to be patterned in this manner, we may see our way to a possible explanation not only of the after-excitation phenomenon but of several others as well.

The foregoing representation of intracortical relationships can hardly be considered imaginary. When we consider the histological structure of the cortex as set forth in standard treatises of neural anatomy (*cf.* Quain, v. 3, 365–366) and note the so-called *plexiform layers* with their *horizontal cells and processes* and the cells of Martinotti, we are persuaded more than ever that the schemata which have been presented above bear more than a superficial resemblance to fact.

In the accompanying diagram, *m* represents an effector; *a, b,*

c, u, d, e, f, represent centripetal fibers conducting impulses from receptors. Now it is clear that if a sufficiently strong excitation advances over *u,* let us say, there will be a diffusion of the excitation over some portions of the network. It is further evident from the nature of the neurone relationships indicated in the diagram, that the wave of excitation caused by the stimulus does not confine itself to advancing radially outward on all sides after the fashion of light or sound waves, but *returns upon itself,* as it were, through a variety of paths. In other words, the excitation, by virtue of the chemical nature of the nerve impulse is *self-sustaining for a period of time.* This is true regardless as to whether or not the excitation " breaks through " to the efferent projection center and fibers.

Should the excitation break through and reach the efferent fiber, however, it is clear that the *frequency* and *duration* of the impulse train passing over the synapse leading to that fiber is determined and conditioned by the entire nerve mass which is participating in the diffusion process. Furthermore, if the excitation process aroused in the cortical segment in question is sufficiently intense, it should be expected that the duration of the response should outlast the duration of the stimulus by virtue of this after-excitation process.

This lack of coincidence between the duration of the stimulus and the duration of the response is a commonplace even in decerebrated animals, and may find its explanation in some such process as the one described. The preceding discussion would lead us to conclude, furthermore, that this discrepancy becomes greater, the greater the amount of internuncial tissue present mediating the excitation and after-excitation process.

We may now proceed to a consideration of (1) the change in the intensity of the after-excitation with time and (2) the change in the intensity of the excitation with increasing distance from its locus of origin (that is, the point where the centripetal impulses enter the cortex).

(1) If we accept the neural mechanism underlying the phenomenon of after-excitation which has been suggested, we may readily deduce from the considerations set forth in Section I

that the intensity of the after-excitation at any given point in the cerebral segment must diminish with time. We may expect this from the adaptation factor involved in our fundamental equation. The relation between the intensity of the after-excitation due to any given stimulus and the time elapsed since its inception we may believe to be some form of an inverse negatively accelerated function. (However, *cf.* Section I.) A quantitative expression of the relationship between the intensity (frequency of impulse) at a given synapse and the time elapsed since its inception is further complicated, of course, by the effect exerted by the network-like structure of the cortical mass. Examination of Figure 2 reveals that the possible summation effects at any given synapse is determined by the conditions at many synapses in the network. These selfsame conditions must necessarily bring about Wedensky inhibitions at some of the synapses and thus institute a local change which resounds, as it were, throughout the whole pattern.

(2) When we come to consider the intensity of the diffusion process with increasing distance from its locus of origin, we must bear in mind Postulate 5, in accordance with which $f_{xa} > f_{yd}$ by virtue of the greater degree of polarization at the synaptic junctions. It might, however, be urged that the irritability of some dendrites might be so much greater than that of the end-brush of their juxtaposed neurone that, in spite of the increased polarization at the junction, the dendrites in neurone *y* might respond with a greater frequency than the frequency transmitted by the axone of neurone *x*. This would mean that a given single impulse transmitted by the axone of the nerve cell *x* would give rise *to more than one impulse* in the dendrites of the juxtaposed neurone *y*. This, however, would seem to be rather improbable because of the refractory phase of the first impulse instituted in *y*, which makes this latter relatively insensitive to further stimulation until it has recovered. At all events, it is a well known fact that a current of constant strength when applied to a nerve fiber will give rise usually to but one impulse, for, as Adrian (p. 63) remarks, "unless the current is very strong

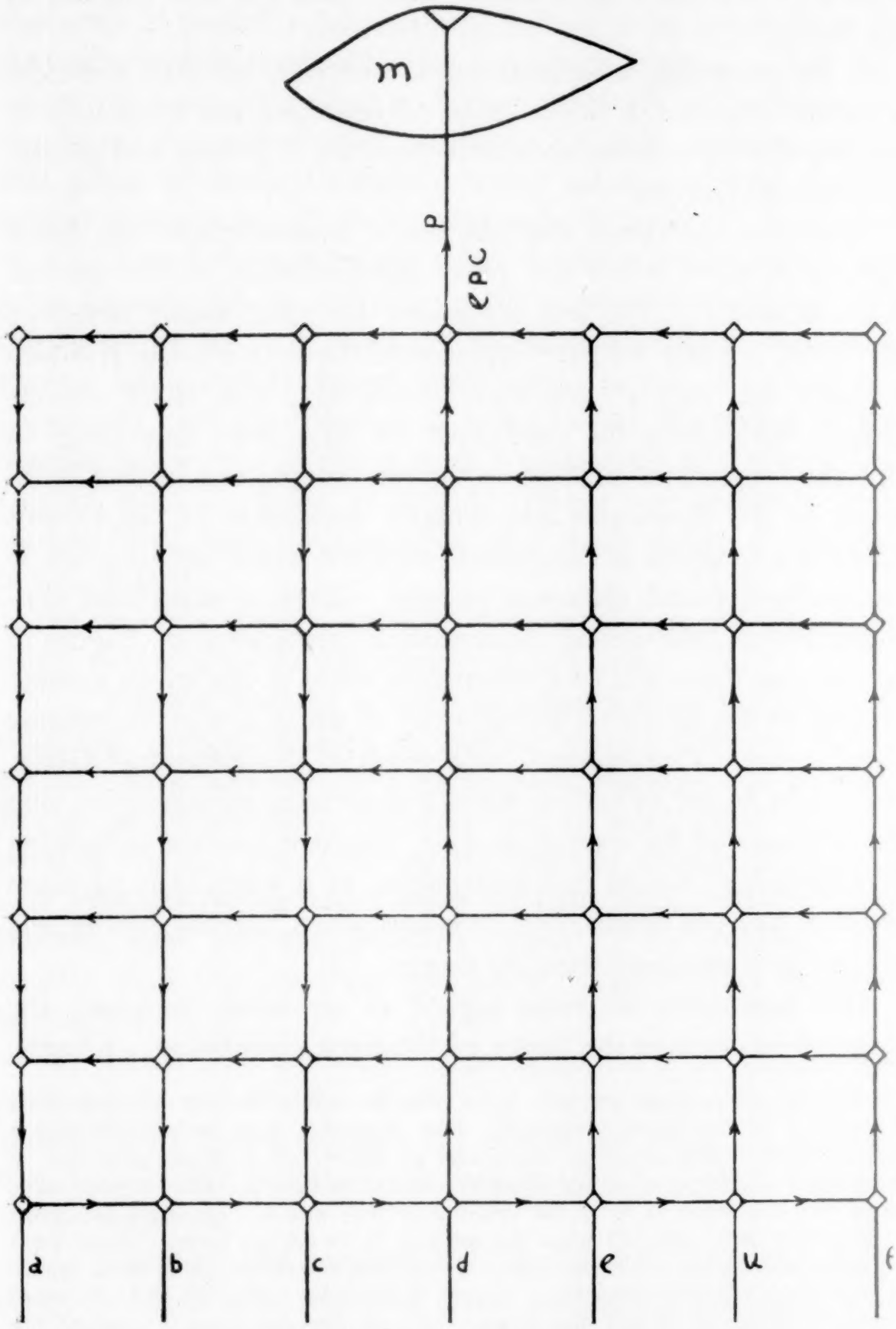

FIG. 2. Schematic diagram of neural relationships. Unidirectional conduction at the synapse is assumed. *m* represents an effector, and *p* the final common path. *u* is the unconditioned stimulus fiber; *a, b, c, d, e, f* are conditioned stimulus fibers. *e p c* is the efferent projection center.

indeed, it will never give rise to more than the first impulse of the series."

If the preceding considerations are correct, we may therefore conclude that, for a single chain of neurones, the frequency of the impulse train diminishes as it traverses a greater and greater number of synapses.*

It is clear, however, that the above considerations are based upon conditions which are never approximated in the cortical mass in reality. The real conditions are more nearly described in Figure 2 where we have represented the network-like relations between the various neurones. Evidently, this latter characteristic complicates the conditions we have been describing, in that the frequencies striking any given synapse are further influenced by the chronaxies and synaptic resistances of the various fibers and synapses in the cortical segment which are involved in the excitation and diffusion process. Even a superficial consideration of the neural relationships illustrated in Figure 2 reveals that there will be a summation effect at any given synapse because of the nature of the network of which any given synapse is but a part. Consequently, as a result of this additional factor, one might be led to believe from a superficial consideration that the intensity of the excitation (*i.e.,* frequency of the advancing impulse trains) would gain momentum, as it were, as it advances from its locus of origin, and, in consequence, become greater and greater as it proceeds from its source.

This conclusion we must regard as erroneous, however, the moment we consider the factor of Wedensky inhibition. Clearly,

* We are,—we must confess it,—a bit charitable in our treatment of Postulate 5 in the above paragraph. For, evidently, it is conceivable that a single impulse transmitted by the axone of nerve cell x might give rise to more than one impulse in the dendrites of nerve cell y. This would occur when (1) the value of θ_o at the synapse is very small, (2) when the value of ϕ_1 is very large, and (3) when the value of f_{xn} is not too large. Under these combined conditions, we might obtain a situation where the value of ϕ_1 would be sufficiently large to institute a second impulse in y after it had recovered from its refractory period and before the next impulse from x reached the synapse. However, since the probability of such conditions obtaining at synapses is smaller than the probability of their not obtaining, it will be seen that the net statistical result for thousands of synapses will be in the direction which we have indicated.

the summation effect at any given synapse due to the network character of neural relationship is limited at its upper limit by the overcrowding threshold at the synapse in question. Not only cannot the summation effect go beyond this limit; on the contrary, should the summation effect transgress the limit of frequency at that point, Wedensky inhibition would set in and thus prevent further transmission of the impulse train altogether. This effect would resound throughout the entire dynamic field involved, and bring about a general lowering of the impulse frequencies. We can therefore see that, since the summation effect at any given synapse has a sharply defined upper limit, the only thing that can happen, in view of Postulate 5, is a gradual diminution of the impulse frequencies as the excitation process diffuses farther and farther from its locus of origin. Hence our preceding conclusion based upon considerations of a single chain of neurones is not vitiated.

We may now concern ourselves with certain general deductions which follow from the preceding considerations.

(1) When a train of afferent impulses reaches the brain the excitation process diffuses throughout certain portions of the cortical mass, and beyond a given point becomes less and less intense as the excitation advances farther and farther from its point of origin. That is to say, the resulting frequencies passing over synapses having many more synaptic junctions intervening between them and the point of origin of the excitation are smaller than the resulting frequencies passing over synapses having less.

(2) There will be certain outlying portions of the cortical mass which will not be affected by the excitation in question because the diffusion process has become so feeble by the time they are reached that their absolute threshold of excitation will not be surpassed.

(3) Applying our fundamental formula of synaptic modifiability, it will be seen that as a result of frequent applications of the stimulus in question, the synapses located near the locus origin of the excitation will benefit in terms of lowered "synaptic resistance" more than those farther removed from the locus

origin of the excitation. There will be points in the topography of the diffusion process where synapses, as a result of the frequent occurrence of the excitation, will suffer from *increased* synaptic resistance (see fundamental synaptic modifiability equation). The results of the process may be seen in Figure 3, where the dynamic relationships are shown in two dimensions only.

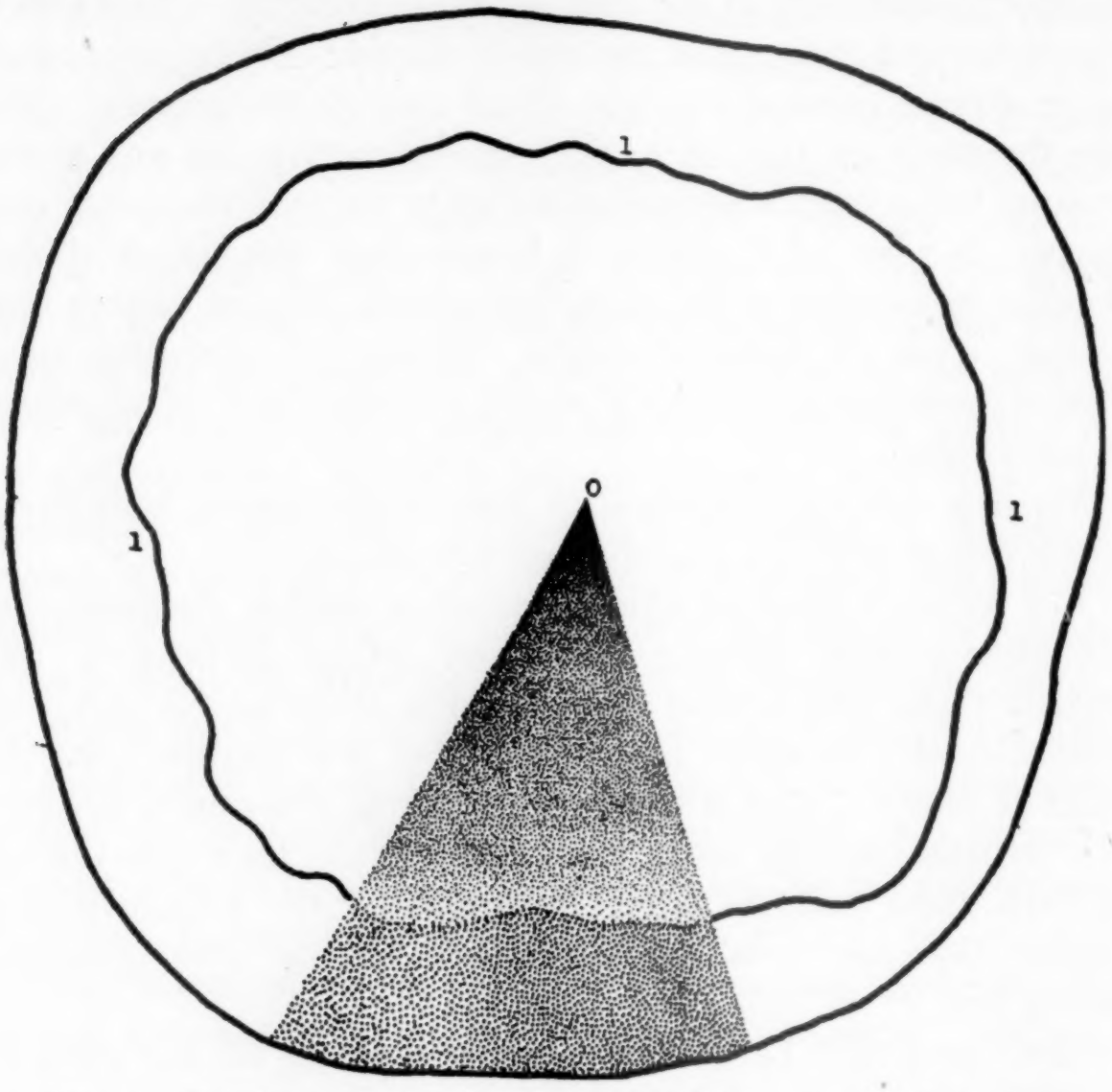

FIG. 3. Density represents synaptic irritability. *o* represents locus-origin of excitation; *l*, the peripheral limits of excitation. The state represented in the diagram occurs only after the stimulus responsible for the excitation has been applied several times.

Figure 3, however, is a representation of the topography of synaptic irritability in some synaptic zone for any given stimulus after a great many applications of the stimulus have occurred. In other words, the diagram represents the end-state. We have interest, in addition, in following the process in its course.

Since the zones in the immediate vicinity of the locus origin

of the cerebral excitation carry a more intense excitation process than the zones farther and farther away (see preceding discussion) it follows, in light of our fundamental modifiability equation, that the amount of increment in impulse frequency at the synapses in the vicinity of the locus origin will be positive and larger than the values at synapses farther removed. Since a positive $\triangle f_s$ value represents a diminished " synaptic resistance " (see discussion in Section I), it must follow that with each successive application of the stimulus, the impulse frequencies passing over the synapses situated near the locus-origin of the excitation must increase in a sigmoid manner. (See Section I.) This in turn must lead us to believe that the area (or, more correctly, volume) of diffusion of the excitation process must *spread* or increase with each application of the stimulus. This, of course, does not go on indefinitely, because the modifiability function approaches a limit. At some point in the process of repetition the increments of synaptic irritability even at the immediate vicinity of the locus-origin approach zero as a limit and therefore become negligible. When this point is reached, there can be no further expansion of the diffusion process.

Since at this stage of the process, repeated further applications of the stimulus in question cannot result in further expansion of the area of excitation, it must follow that the periphery of this area, beginning at the extreme edge, and proceeding inward for a certain distance, will begin to suffer from increased polarization. That is to say, the values of f_s at the periphery will decrease, the largest decrease being at the outermost edge of the diffusion zone. It may be seen from these considerations that continued repeated applications of the stimulus will diminish synaptic irritability here and produce in effect a gradual constriction of the field of excitation caused by the stimulus. This constriction will continue until the $-\triangle f_s$ values approach zero, when further applications of the stimulus effects no further change.

We thus see that repeated applications of a stimulus give rise to an initial expansion of the area of excitation, followed by a small gradual constriction, until a certain equilibrium is reached.

In the preceding discussion of the diffusion process we have

confined our attention to the sensory areas of the cortex, and have proceeded as though these had no connection with the frontal lobes and their projection centers. This, of course, is patently not true, for we have at least two bundles of association fibers which directly connect the sensory and motor areas in a given hemisphere, in addition to the transverse fibers constituting the corpus callosum, anterior commissure and fornix, which connect the two hemispheres. The former two bundles are the *fasciculus longitudinalis inferior* and the *fasciculus uncinatus.* Taking just one hemisphere of the cortex, therefore, and exploring its surface, we should find, upon stimulation of some receptor, an area of increased potential at the appropriate sensory projection centers and another area of increased potential in the frontal lobes, this latter area having been made possible by the conduction of the longitudinal association fiber bundles. The distribution of potential on the cortex under the above conditions, therefore, would give a sort of dumb-bell effect. An analogous condition would of course be found in the other hemispheres, owing to the conduction of the transverse commissural fibers. This dumb-bell appearance would be more or less pronounced in accordance with whether the sensory and motor projection areas involved were more or less widely separated.*

The conclusions reached in previous paragraphs, therefore, in regard to the behavior of the area of excitation in the sensory cortex with repeated applications of the stimulus, hold with equal justification for the area of potential in the frontal lobes. The discussion, consequently, presents certain implications which are of considerable theoretical importance.

Clearly, if we imagine our area of excitation due to the stimulus to include various efferent projection centers, it must follow that on the first few applications of the stimulus some of these centers will be excited and will produce diffuse activity in

* Several months after this was written, and as the manuscript was being sent to the press, the paper by Perkins on *Cerebral Action Currents in the Dog* became available to the writer. The results reported in that paper together with the topographical maps of the electrical potential gradients in the dog's cortex give ample corroboration to the considerations which have been advanced in the preceding pages. (*Psychol. Monog.,* XLIV, 1, 1933.)

the appropriate effectors. It should follow from our discussion that the amount of diffuse activity should increase with the first few applications of the stimulus (if the stimulus is really new to the organism), as the area of excitation increases, but that at some point an inversion should take place, as a result of which the diffuseness of the reaction should gradually diminish until it reaches some point which is a limit. The latter portion of this deduction, at any rate, we know to be corroborated by fact.

III

Forward Conditioning

The phenomenon of the conditioned response, systematically investigated by Pavlov and his students, may be considered as one of the simpler manifestations of the learning process. We need not dwell here upon the details of the procedure requisite to obtain simple conditioning, as these have become common psychological knowledge. We need only to mention the fact that it is customary, in instituting the process of conditioning, to present the conditioned stimulus a second or so before, or simultaneously with, the unconditioned stimulus. Whether or not this is an indispensable condition for the phenomenon, we shall presently consider.

We may conduct our discussion in terms of the schematic diagram shown in Figure 2. We shall choose *u* as our unconditioned stimulus fiber, and *a* as our conditioned stimulus fiber.

Now, it is clear from our previous discussion that if the conditioned stimulus is applied at *a,* there will be a diffusion of excitation throughout a portion of the cortical network, subject to the conditions peculiar to such an event. By definition, the intensity of the diffusion process is not sufficient to break through to the final junctions and the efferent path; otherwise an unconditioned response would occur. Now, as we have seen, the excitation does not die out as soon as the conditioned stimulus is removed, but remains in the form of an after-excitation whose intensity, however, does diminish with time. Presently, the unconditioned stimulus is applied, and produces its characteristic diffusion of excitation which shortly reaches the efferent projection center, among other things, and produces an unconditioned response.

Now, it is clear that if the excitation from one stimulus is to be modified by the excitation from the other stimulus, the areas of excitation aroused by the two stimuli must to some degree overlap. If there is no overlapping, then presumably no modification

can take place, as there would be two independent dynamic systems with no means of influencing each other.

We are now ready to examine the nature of the mutual modification resulting as a consequence of the overlapping of the excitation due to the unconditioned stimulus and the after-excitation remaining from the conditioned stimulus. It is clear that neither the conditioned stimulus alone, during the course of the diffusion of the excitation caused by it, nor the after-excitation surviving it, was able to achieve the result of evoking the response in question. We must now suppose that, since the two excitations overlap, there is a mutual reinforcement, such that a resultant excitation of unprecedented intensity sweeps over the final junctions to the efferent projection center. In other words, whereas, when the conditioned stimulus is applied alone, a zone of too high resistance prevented further diffusion of the excitation process in the crucial direction of the projection center to the final common path, the reinforcement or intensification in the overlapping zone brought about by the diffusion of the unconditioned excitation has now caused a new wave of excitation of greatly increased intensity to transgress this crucial area.

This would lead us to conclude that the unconditioned response elicited by the unconditioned stimulus, coming as it does, in the wake of the conditioned stimulus, should be somewhat more vigorous than when the unconditioned stimulus is acting alone. This has, in fact, been found to be the case in Hilgard's (16) careful investigation. It is also apparent that if the intensity of the unconditioned stimulus is optimum, the addition of a conditioned stimulus preceding it in time may effect a decrement of impulse frequency at the efferent projection center owing to the Wedensky effect, and thus bring about a diminution in the vigor of the response.

We must now be reminded of our discussions in Section I. It will be remembered that we concluded there that the change in polarization at the synapse, θ being given a value, is a function of the frequency of the impulse train which has passed through it. This is more fully expressed in equations (3) and (8). We may see, therefore, that, with repetition of the combined conditioned and unconditioned stimuli, the synaptic "resistances"

in the zone which originally impeded the excitation of the conditioned stimulus from flowing to the projection center and stimulating the final common paths, may be reduced; hence, eventually, the application of the conditioned stimulus alone will evoke the response.

Before going further, it is well to investigate a little more closely how it is possible for reinforcement to take place in the area of overlapping. That reinforcement should take place is not at all obvious, for although it is generally known that increased stimulation results in a greater frequency of impulses in the train, it may also result in these particular conditions in partial or complete inhibition.

It is clear that the zone of overlapping of the unconditioned and conditioned excitations must involve many fibers, and hence many paths and combinations of paths. In some of these paths the overlapping of excitations, because of Wedensky effects, will result in a diminution of the impulse frequency, and in others, in a total abolition of excitation; in the remainder it will result in an increase in frequency. Conditioning will therefore take place along those paths and combinations of paths where the frequencies have been increased and the synaptic irritabilities have been therefore sufficiently heightened by the various frequency values to permit of the flow of the conditioned stimulus excitation to the projection center controlling the response to be conditioned. Were these crucial segments in the cortical mass destroyed, presumably the conditioning previously established would disappear.

Our equations would lead us to believe that the process of conditioning will be sigmoid if the overlapping areas are sufficiently "pristine"; otherwise it will be negatively accelerated.*

A number of interesting phenomena may be discussed in light of the above considerations. According to Postulate 2, the intensity of the after-excitation to any stimulus varies as some inverse function of the time. Hence, there are certain temporal

* Should the reader wonder why we assume implicitly that a greater and greater impulse frequency bombarding the efferent projection center should result in an increasingly vigorous response, he need only reflect that the higher the frequency reaching the projection center, the greater the number of efferent fibers which will probably be activated.

limits within which the conditioned and unconditioned stimuli must fall before the simple conditioned response phenomenon may be instituted. As the temporal interval increases, the intensity of the after-excitation of the prior stimulus decreases, and the amount of maximum reinforcement which may take place is accordingly less, the limit being zero when no conditioning may take place. Within this limit, however, it must follow from our presuppositions that the greater the temporal interval between the two stimulations, the greater the number of repetitions required to produce conditioning. Furthermore, it is presumable that the function expressing the relationship between the number of repetitions required to establish simple conditioning and the temporal interval separating the conditioned and unconditioned stimuli should somewhat resemble the function expressing the relationship between the intensity of the after-excitation at a given locus in the cortical mass and the time since its inception.

It further follows that, if a temporal series of conditioned stimuli, *a, b, c, d, e* precedes the unconditioned stimulus, the conditioning will first be set up with reference to *e,* then proceed backward in the order *d, c, b, a.*

Our considerations would further lead us to believe that, once conditioning has been established, an increase in the intensity of the conditioned stimulus (within a limit) should increase the intensity of the conditioned response.

It might also be expected from our discussion of the conditioning process that the application of a *sufficiently intense* unconditioned stimulus should so diminish the synaptic resistances in the crucial overlapping zone so as to make it possible for the excitation due to a heretofore innocuous stimulus, whose radius of diffusion is sufficiently large, to gain access to the projection center and therefore to the efferent tract. This apparently is what actually happens in abnormal or neurotic cases. Nearly every animal investigator has had occasion to observe cases where, through some unfortunate accident, a particular animal, having received some violent sensory shock, is subsequently rendered useless for further experimental work. Any sound, movement, light or touch—in fact, any stimulus whatsoever—is likely to cause the animal to jump or cringe.

IV

Backward Conditioning

We have thus far confined our remarks to the phenomena known technically as "forward" and "simultaneous" conditioning. We must now consider "backward" conditioning. Here we have the conditioned stimulus *following* the unconditioned stimulus—a fact which obviously alters somewhat the situation.

In this case, the unconditioned stimulus, by virtue of the native structure of the nervous system, has set up a train of impulses which, single handed, has evoked the appropriate response. We have, however, the phenomenon of after-excitation, possessing properties of distribution and intensity which bear a fundamental relationship to the original pattern of potential which constituted the diffused excitation. Thus the train of impulses instituted by the conditioned stimulus a few moments later does not come, to speak figuratively, upon neutral ground. The response has occurred, however, by this time, and unless the intensity of the persisting after-excitation due to the unconditioned stimulus is such that, combined with the excitation of the present conditioned stimulus, it breaks through to the efferent center, nothing observable will occur. If, however, the combined excitations do "break through," we have a repetition or continuation of the response—probably of lesser intensity—and hence a basis for simultaneous conditioning. Sometimes this subsequent breaking through does occur; this is shown where the organism is subjected to some powerful stimulus, such as an electric shock, which elicits an immediate response. If shortly thereafter the organism is stimulated with an ordinary innocuous stimulus, such as a light or a sound, the response is likely to be given.

As has been indicated already, if such reaction occurs, and is successfully repeated several times, we may succeed in instituting a conditioning process which, actually, is of the "simultaneous" kind.

There has been, however, in this country at least one recent investigation (Switzer's, 27) in which the conditioning process is instituted by presenting the unconditioned stimulus first, even though there was not a breaking through of the second stimulus (the conditioned stimulus) upon its first presentation as described above.

In Switzer's experiment we undoubtedly have a case of true backward conditioning:—in the nature of an unconditioned response to the native stimulus, followed by a feebler response to the conditioned stimulus, *after the training had been completed.*

It is necessary to look again into the possible neural processes which we assumed to underlie the conditioning process. We have seen that conditioning is made possible by the overlapping of the excitation zones due to the conditioned and unconditioned stimuli. This makes possible a reënforcement of excitation which, in turn, sufficiently increases synaptic irritability to permit the conditioned excitation alone to break through in the region originally hampering its advance toward the projection center.

Now, on the basis of the considerations in regard to the diffusion of excitation which we advanced in Section II, it is evident that the degree of reënforcement consequent upon the overlapping of the excitations due to the conditioned and unconditioned stimuli is definitely influenced by the temporal sequence of application of these stimuli. Inspection of Figure 4 reveals that the sequence conditioned stimulus–unconditioned stimulus gives a greater reënforcement effect with reference to the efferent projection center in question than the sequence unconditioned stimulus–conditioned stimulus. We may therefore see from the logic of the situation that backward conditioning is at a distinct disadvantage, and that a much greater number of repetitions is required than in forward conditioning. In fact, the amount of reënforcement may be so small that the conditioning process is impossible.

A glance at the schematic diagram shown in Figure 4 would lead us to infer that efferent projection centers functionally favored by the conditioned excitation would be likely to be affected much more by the backward conditioning process than by the forward conditioning process. We may illustrate this by

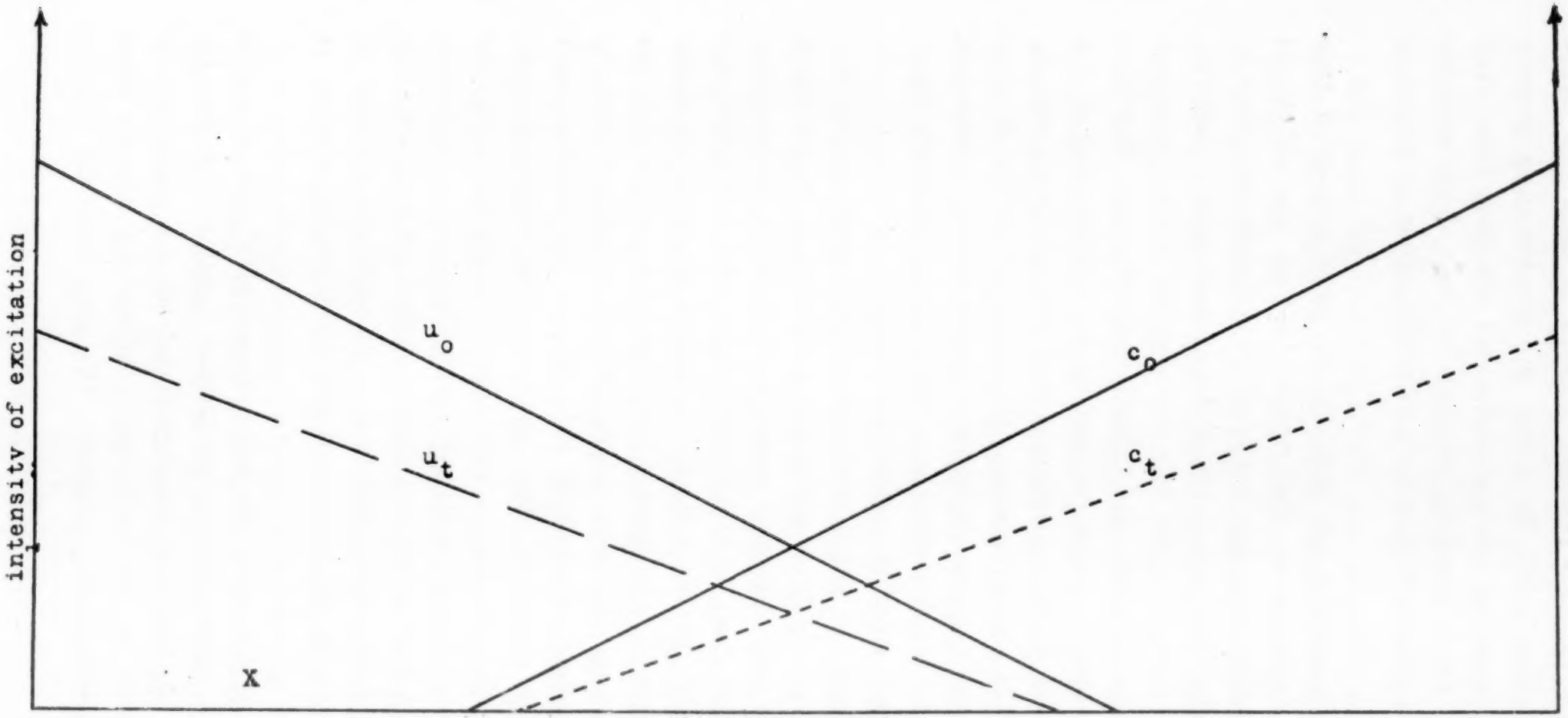

FIG. 4. x represents the locus of projection center of the response to be conditioned. u_o represents level of potential of the unconditioned excitation immediately after application of the unconditioned stimulus; u_t, the level of potential of the unconditioned excitation when the conditioned stimulus is applied, if the sequence is unconditioned-conditioned stimulus.

c_o represents the level of potential of the conditioned excitation immediately after the application of the conditioned stimulus; c_t, the level of potential of the conditioned excitation when the unconditioned stimulus is applied, if the sequence is conditioned-unconditioned stimulus.

For the sake of simplicity, the overlapping effect for the two kinds of sequences is not shown. The reader may supply this for himself, and may note that, given the particular relationship of the efferent projection center to the two excitations, the sequence c.s.-u.s. is the more favorable to conditioning.

an example. Let us suppose that in a backward conditioning experiment the conditioned stimulus (which is presented second in the series, of course) is a fairly loud noise which produces twitching of the ears. The unconditioned stimulus is an electric shock producing retraction of the foot. We might expect that in such a case, the twitching of the ears would be more likely to remain and to become more pronounced in backward conditioning than in forward conditioning when the same stimuli were used.

Our preceding considerations enable us to understand further, how, under such conditions, the electric shock alone would produce twitching of the ears.

We may concern ourselves at this point with the frequently attested fact that the latent time of response to the conditioned stimulus is distinctly greater than the latent time of the same response to the unconditioned stimulus. We know from recent evidence that in the intact animal the stimulus leading to even the lowest type of reflex response brings about the after-excitation phenomenon in the cerebral mass. We know, furthermore, that these elementary reflexes may be elicited in spinal animals, hence in organisms where any progress of the impulse into the brain mass is impossible. On the other hand, it has likewise been amply demonstrated that a conditioning process cannot be set up when there has been excessive destruction of the cerebral cortex. It is thus seen that the process of modifiability which we have been describing demands as one of its conditions a more or less intact cortical mass. This can only mean that the excitation from the conditioned stimulus alone, in producing a conditioned response, must traverse a longer path and overcome a greater amount of neural inertia before it reaches the effector; hence an increase in latent time is quite in order. A similar relationship is observed when we compare the latent time of reflexes and the so-called voluntary latent time—and for the same reason.

The process underlying the phenomenon of modifiability in the higher forms of animal life, then, seem to involve invariably the reticular cortical mass. This perhaps may account for the fact that Lashley was unable to discover any evidence that, with continued repetition of the learned act, the neural processes involved tended to be relegated to the lower centers.

V

Experimental Extinction

We may now move on to a consideration of experimental extinction. It is evident at the outset that this is not to be confused with the phenomenon of inert forgetting, which may be adequately conceived as being caused in consequence of repolarization in the absence of the conduction of nerve impulses through synapses. We refer rather to the phenomenon observed when, the conditioning process having been successfully set up, the conditioned stimulus alone is repeatedly applied. As is well known, the conditioning process is rapidly destroyed in consequence:—to quote the Russian school, an inhibition has been produced which makes the occurrence of the conditioned response impossible.

Clearly, experimental extinction is, on the basis of our principles, a phenomenon caused by increased polarization at the synapse due to the transmission of inadequate frequencies. Thus while the combined excitations due to the conditioned and unconditioned stimuli served to produce frequencies sufficiently high to effect considerable depolarization at the synapses and thus to permit the lower frequencies of the conditioned stimulus excitation to break through and reach the appropriate efferent projection center, it is evident that these lower frequencies, by continuing, increase the amount of positive charge at the synapses. (See equations in Section I.) We have seen that a process of this kind, when repeated, diminishes synaptic conductivity at a negatively accelerated rate (p. 15). Hence the conditioned response, when not reënforced now and then by the application of the unconditioned stimulus, will diminish in vigor at a negatively accelerated rate, as the conditioned stimulus continues to be applied alone.

A comparison of this deduction with the experimental results reported by Pavlov (**24**, p. 52) shows good agreement between them.

It is evident that experimental extinction need not necessarily follow as a result of the repeated application of the conditioned stimulus alone—provided the interval between successive applications of the unreënforced conditioned stimulus is not too short. (*Cf.* Section I, p. 18.) Whether experimental extinction will occur under these conditions will be determined by the values of θ at the various synapses in the critical overlapping region and by the functional relationship existing between the locus origin of the conditioned excitation, the critical overlapping region and the efferent projection center. If the functional relationships between these three regions are "good," the f_s values sent through the synapses in the overlapping region by the conditioned stimulus excitation alone may be sufficiently high to make for a steady maintenance of the "connection," or even for a greater degree of fixation, as a result of repeated application of the conditioned stimulus.

Hence it is not at all at variance with our principles to assert that a conditioned reaction may last for the course of a lifetime without further reënforcement by the unconditioned stimulus. Whether this happens or not is, as we have indicated already, determined by the functional relationship between the cortical regions involved, plus the nature of the unconditioned stimulus, the unconditioned response and the conditioned stimulus.

It is interesting to note in this connection that Pavlov and his coworkers report abundant cases of a resumption after twenty-four hours of the conditioned response subsequent to its extinction. On the basis of the present theoretical discussion, we should expect this to occur if the temporal interval separating the successive applications of the unreënforced conditioned stimulus is not too long. For if the intervals separating successive volleys of impulses in the cortical mass are not sufficiently long to permit of complete recovery from synaptic adaptation, there may be a cumulative adaptation effect which serves not only to increase the amount of positive charge at the synapse (Section I) but to so diminish the irritability of the nerve tissue there as to prevent, at some point in the process, the further conduction of impulses through the synapse. When this occurs the con-

ditioned response disappears. If, now, the conditioned stimulus is not applied for some time, the adaptation will have become dissipated to a sufficient degree and hence an application of the conditioned stimulus a few hours later will evoke the response. This will be true, of course, only when the process of experimental extinction has not brought with it a repolarization of the synapses sufficiently great to block the conduction of the conditioned stimulus excitation; that is to say, the phenomenon will occur only when the synaptic adaptation process has succeeded in impeding the passage of impulses across the synapse before they succeed, with their low frequencies, in critically repolarizing it. For, according to our considerations, while synapses may recover spontaneously and somewhat completely from adaptation, they do not recover from polarization to the same degree. (Polarization recovery is made possible by diffusion.)

From the discussion in the preceding paragraph, it is manifest that there is much less likelihood of there being spontaneous recovery of the conditioned response when the unreënforced conditioned stimulus, during experimental extinction, is applied at relatively great intervals as when it is applied at short intervals.

Two further deductions are implied in our considerations:

(1) The extinction of a conditioned response occurs sooner if the applications of the unreënforced conditioned stimulus be given in quick succession than if they are separated by a large temporal interval. As is well known, Pavlov finds this to be true.

(2) If conditioning has taken place to a stimulus of a given intensity, and then this stimulus is applied without the reënforcement of the unconditioned stimulus, the number of repetitions required to institute experimental extinction, everything else remaining constant, will vary as some direct function of the intensity of the conditioned stimulus.

Finally, let us consider, in the light of the preceding considerations, the following hypothetical situation:

A stimulus has been applied numerous times to the organism at constant intervals for a constant period of time and has been evoking a response whose intensity has been increasing, let us say, at a negatively accelerated rate. Once the limit of the response

intensity is reached and maintained for some time, we cease applying the stimulus. After an interval of inactivity of a few hours, we apply the stimulus once more and obtain a response greater than any obtained heretofore. What has happened?

We may see our way to an explanation of this phenomenon if we consider the interval separating the successive applications as being such as to permit a degree of synaptic recovery which, during the plateau, merely balances the effect of increased *de*-polarization instituted by the preceding volley. When the succeeding volley comes, the synapse is in the same condition as when the preceding volley began. This must hold true for each successive interval, hence there will obtain a cumulative synaptic adaptation effect which, at the moment of application of the stimulus, is just balanced by the cumulative *de*-polarization effect. This must mean that, during the plateau, while the value of θ is found diminished at the beginning of each successive impulse volley the value of A is found increased by a corresponding amount, so that the resultant f_s does not change from one application of the stimulus to the next.

If, now, an interval of nonstimulation intervenes, there will be opportunity for adaptation recovery at the synapses, and when the stimulus is applied once more, the diminished θ values at the synapses, heretofore masked by the adaptation process, make themselves apparent by an increase in the vigor of the response. This can be expected to occur only when the period of non-stimulation is not too long; otherwise the repolarization process, slow as it is, will mask the effect.

Perhaps such a mechanism as this underlies the phenomenon described by James as " learning to skate in summer and to swim in winter."

VI

The Conditioned Trace Response

The " trace reflex " phenomenon, as it is called by the Russian investigators, is interesting by virtue of the fact that it seems to contradict the essential conditions found necessary in establishing the ordinary conditioned response. The phenomenon occurs under the following conditions:

The conditioned stimulus precedes the unconditioned stimulus as in ordinary conditioning set-ups; the one difference being that the application of the conditioned stimulus continues until the unconditioned stimulus is applied and the temporal interval separating the initiation of the two stimuli is greatly increased.* Now, if the pair of stimuli are repeated in this temporal sequence, eventually the conditioned response will appear in answer to the conditioned stimulus, but only after an interval of time roughly approximating the interval separating the *beginning* of the conditioned stimulus and the application of the unconditioned stimulus. It would appear, then, that in evoking the trace response, the conditioned stimulus alone is not sufficient to determine the reaction. We must have, in addition, its *continuation* for the appropriate period of time.

In attempting a possible neurological explanation of this phenomenon we must keep in mind the duration aspect of the cerebral excitation due to the prolonged application of the conditioned stimulus. During the training process, the conditioned stimulus is applied, and an excitation of a given locus and intensity is set up. The extent and intensity of this excitation varies, for one thing, with time, so that the moment of the arrival of the unconditioned stimulus finds the process at a stage of intensity and spread determined by the time elapsed since its inception. When

* The considerations to be employed here may be entertained in discussing a situation where the duration of the conditioned stimulus does not overlap the unconditioned stimulus.

the unconditioned stimulus arrives, there is a diffusion of cerebral excitation and a consequent overlapping of excitation zones with the concomitant reënforcement effect in the area of overlapping. This, as we have seen, serves to lower the synaptic resistances of the paths and combinations of paths leading to the projection center.

Evidently, thus far we have considered the typical conditions for the setting up of simple conditioned response; it is clear, however, that we have omitted some essential factor, because, even after many repetitions, the conditioned response is not evoked immediately upon the application of the stimulus: there is a delay corresponding in amount to the size of the temporal interval originally employed.

Our cue lies in the fact that the prolonged application of the conditioned stimulus increases the value of the t variable in equation (8) and that, in consequence, at the moment of arrival of the unconditioned stimulus, the spread and intensity of the field of excitation due to the conditioned stimulus is not as great and as intense as it would be had the conditioned stimulus been applied only briefly and the unconditioned stimulus been given immediately thereafter.

In consequence of this fact, the amount of reënforcement in the area of overlapping will be less and the decrement in synaptic resistance in that area and in immediately surrounding areas correspondingly smaller. This accounts for the fact that a greater number of repetitions are required to institute a delayed conditioned response than a simple conditioned response.

It is, nevertheless, paradoxical that the conditioned trace response should occur at all. For, in accordance with our principles, the conditioning process is essentially a matter of reducing synaptic polarization so that an excitation may have an effect which previously it did not have. Hence, the response, under the foregoing conditions of the trace reflex, should occur shortly after the conditioned stimulus is applied, for it is then that the intensity of the excitation is the greatest, and therefore may more readily traverse the crucial cortical areas and produce a response. As a matter of actual fact, however, it would seem that, in the phe-

nomenon we are discussing, the response occurs when the intensity of the excitation is (because of the adaptation effect) smaller. It is not evident how this necessary condition of weaker excitation intensity may lead to a response when a condition of greater intensity does not. Yet we are forced to conclude that the weaker excitation is at an advantage in this case, because the continued application of the conditioned stimulus is necessary to elicit the delayed response, and, since we believe that, beyond a certain point, the intensity of the excitation varies as an inverse function of the time elapsed since its inception (see Section I, p. 15), we must be persuaded that the actual intensity of the excitation at the moment the conditioned response does occur is less than at the moment immediately succeeding the beginning of the application of the conditioned stimulus.

Our problem, then, is to determine the conditions under which lowered synaptic resistances in certain zones will permit a weaker rather than a stronger excitation to break through and evoke a response. (We must admit, on the basis of our principles, that a state of lowered synaptic resistances or polarization in the crucial zones has been achieved by the conditioning process, otherwise the conditioned stimulus would not release the response.)

At this juncture we are reminded of the Wedensky inhibition phenomenon. We have, in this neurologically attested fact, a situation which corresponds quite strikingly to the more complicated situation we are considering:—an impulse train of smaller frequency evokes a response where an impulse of greater frequency fails: a matter of the succeeding impulses falling into the refractory phase of the preceding one. If we keep this phenomenon of overcrowding in mind we may see our way to a plausible solution. It is reasonable to assume that the myriad neurones and synapses in the cortex have varying overcrowding limens. Let us suppose a set of neurones and synapses that have a rather low overcrowding limen. When a stimulus of a given intensity is applied to the organism, the resulting nerve frequencies aroused in the cortex are such that Wedensky inhibition sets in for these particular sets of fibers and synapses. This means that no impulses will be transmitted over them until the

extant frequencies in the cortex are sufficiently lowered. If, however, the conditioned stimulus continues to be applied for some time the adaptation effect will lower these extant frequencies. When the frequencies have been sufficiently lowered, they will be transmitted over these hitherto impervious paths.* Let us assume, now, that at this point the unconditioned stimulus is applied. As a consequence, we shall have a reënforcement effect. This must be small because if the resultant frequency is too high it will again bring about Wedensky inhibition. In fact, there are undoubtedly a great many synapses and fibers suffering from Wedensky inhibition because of the reënforcement effect due to the unconditioned stimulus. However, we may believe that there will be some of these fibers having low frequency limits, which by virtue of the reënforcement due to the unconditioned stimulus, will conduct to the efferent projection center with slightly increased frequencies. This reënforcement brings about its usual increment in synaptic irritability, which in this case, must be somewhat small because of the lower frequencies involved.† In view of this fact, a greater number of applications of the conditioned and unconditioned stimuli are required to sufficiently lower the synaptic resistances of these low frequency synapses to enable the conditioned stimulus alone to bring about a response. This we know to be a fact.

The preceding discussion would lead us to believe, therefore, that we have in the nervous system a mechanism adequate for the formation of delayed conditioned responses of varying time interval values, although we may be readily persuaded that this range of possible temporal intervals must have definite limits. The value of this range must be a function of the range of the overcrowding limens possessed by the various synapses.

Further consideration of the foregoing processes reveals that the temporal interval separating the inception of the conditioned

* Of course, there may be some low-frequency-limen paths so situated as to conduct immediately after the stimulus is applied, but these will be "occluded" by the adaptation factor as are the others.

† The discussion in this paragraph assumes that the constant b in equation (1) has varying values for different synapses, and that in the case of synapses of the kind under discussion, it will be sufficiently small to make *de*polarization possible. (See Section I.)

stimulus and the conditioned response will tend to become shorter with practice. We must expect this to occur because of the increased irritability imparted to these synapses with low overcrowding limens and the concomitant diminution of positive charges there. This serves to raise the overcrowding limen, for we have seen from Erlanger *et al.* (*loc. cit.*), that weak cathodal polarization serves to decrease the duration of the falling phase of the impulse by approximately 82 per cent. Consequently, excitations of higher and higher frequencies are in position to stimulate the fiber tracts which actuate the conditioned response under the particular conditions. This tendency is further accentuated by the fact that the synapses of fibers other than the critical ones which are immediately actuating the efferent projection centers in the delayed conditioned response become increasingly adapted (up to some limit) with repetition of the conditioned stimulus, and hence the excitation intensity drops at a much more rapid rate. The result of this is that the critical fibers are stimulated sooner—that is, the excitation falls below the overcrowding limen sooner.

The interval of delay between the inception of the conditioned stimulus and the conditioned response cannot become vanishingly small, however, because of two facts: (1) The unconditioned excitation serves to neutralize somewhat the adaptation effect brought about by the prolonged application of the conditioned stimulus in those fiber tracts other than the ones immediately actuating the efferent projection center. (2) If the conditioned response comes too soon before the application of the unconditioned stimulus, the fibers which were responsible for the premature conditioned response become adapted by the weak excitation which permeates them for the period of time until the reënforcement by the unconditioned excitation occurs (since the objective temporal interval between the inception of the conditioned excitation and the unconditioned excitation is kept constant by the experimenter). This state of affairs, therefore, serves in the capacity of experimental extinction.

Two deductions may be made from the preceding discussion:

(1) We should expect that the conditioned delayed response

to a given conditioned stimulus would be less intense than to the same stimulus in the conventional conditioning set-up. This is evident when we consider the fact that in the delayed conditioned response smaller frequencies are involved.

(2) It may further be expected that if, in the trace response set-up, after the delayed conditioned response has been established by means of a conditioned stimulus of a given intensity, we suddenly apply a conditioned stimulus of greater intensity, the temporal interval intervening between the application of the conditioned stimulus and the occurrence of the response would tend to become longer. Conversely, if a weaker stimulus were used, the interval would tend to become shorter.

We may add a word here on the "inhibitory function" of the cerebrum. We have already indicated that the cerebrum is to be considered as an intricate network of fibers which makes it possible for an impulse train which enters the cortex at one point to bring about an excitation which diffuses, returns upon itself (see Section II, Figure 2) through various paths and thus in a most extraordinary way *summates itself*. It is clear that in this way there are times when, through this summation effect, the excitation in any given cortical field may transgress the statistical overcrowding threshold for that field and thus become inhibited until such time as the adaptation process reduces it to limits falling below the Wedencky threshold, whereupon the excitation process advances once more unimpeded to the efferent centers. This factor undoubtedly is influential in retarding the time of behavior reactions which are innervated through the cerebrum. It may also account for the violent and spasmodic reactions of decerebrated animals.

There are only too many times when we are "struck dumb" by a violent sensory situation and are incapable of moving for a time. The possible application of this to the "death feigning" of animals when in the presence of danger is evident.

VII

Conditioned Inhibition

We proceed next to the phenomenon of conditioned inhibition. For the sake of expedience in exposition, the set-up may be described as being essentially the following: A conditioning process has been set up between a given stimulus which we shall call *a* and a given response; conditioning has likewise been set up between another stimulus *b* and the same response. Once this double conditioning has been instituted, one of the stimuli—let us say *a*—is systematically reënforced by the unconditioned stimulus; the other stimulus—*b*—is just as systematically not reënforced. It follows, in consequence of experimental extinction, that stimulus *b* soon loses its capacity to evoke the response. If now we apply both stimuli simultaneously, stimulus *a* fails to evoke the response; or the response is perceptibly weakened. (As a control, in this type of experiment, it is always shown that, before experimental extinction of *b* has been brought about, its presence simultaneously with *a* does not inhibit the latter.)

If we try to envisage the situation neurologically, it is at once clear that the areas of excitation due to stimuli *a* and *b* must in some measure overlap; this must be so, otherwise *b* would be unable to influence *a,* as it actually does. Now, in discussing the phenomenon of simple conditioning it will be remembered that we considered the areas of excitation due to the conditioned and unconditioned stimuli as overlapping to a certain degree, and that with repeated reënforcement by the unconditioned stimulus the area of effective excitation due to the conditioned stimulus was so enhanced as to extend to the efferent projection center governing the response. It is clear in the present situation that the areas of excitation due to the conditioned stimuli *a* and *b* not only overlap, but, more significantly, after the positive conditioning process has been completed for both, overlap in regard to the locus of the efferent projection center.

Past speculations have shown how inadequate it is to think in terms of isolated paths approaching the projection center from separate directions and in such a way as to be unable to influence each other. We must rather think in terms of areas containing hundreds and probably thousands of fibers with all possible relationships to each other. We feel justified, therefore, in using the following figure to facilitate exposition. We are naturally

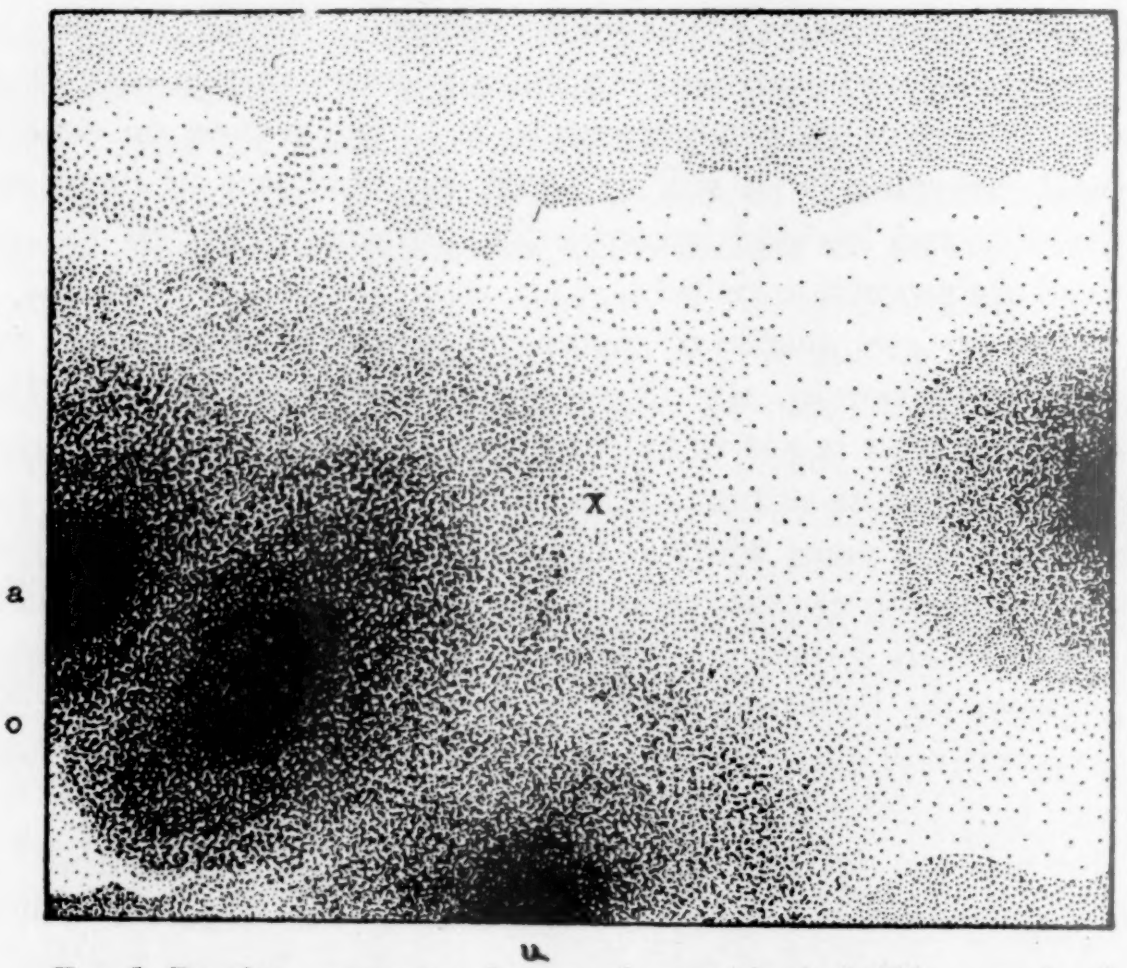

FIG. 5. Density represents degree of synaptic irritability. *a* is the locus-origin of the conditioned excitation; *o*, area of overlapping; *u*, locus-origin of the unconditioned excitation; *b*, the locus origin of the conditioned excitation. X is the locus of the efferent projection center.

unable to express the third dimension in the figure, hence must be contented with a cross-section of the dynamic fields. (Figure 5.)

It will be noted that there are represented in the figure (which represents the state of affairs when the scene is prepared for the occurrence of conditioned inhibition, *i.e.*, conditioning to *b* has been extinguished) four fields of heightened irritability. The one at the extreme left represents the locus of origin of the conditioned excitation *a*; the ones next toward the right, the area of

overlapping of excitation *a* and the unconditioned excitation *u*, respectively; and the one to the extreme right, the origin of the locus of the conditioned excitation *b*. It will be noted that the degree of irritability in the field governed by the conditioned excitation *b* is distinctly lower than in the other fields. This has been brought about by the extinction of the conditioning process to *b*. This process of the extinction of the conditioned response to *b* has brought a lowering of the threshold of overcrowding as well as a lowering of synaptic irritability. In fact, the degree of irritability and overcrowding limen is lower than it was *prior* to the setting up of conditioning to *b*. (See section on experimental extinction.) It will be noted that this area of increased irritability and lowered overcrowding threshold extends to and beyond the projection center.

Since we are unable to represent the third dimension, the impression given by the figure is somewhat erroneous. The projection center is not to be construed as belonging to the same plane as the cross-section of the brain fields represented in the figure, but in a plane which is deeper in the brain mass caudalad. Thus, while the excitation area due to the extinguished conditioned stimulus *b* does not effectively reach the projection center because of the increased resistance brought about by the process of experimental extinction, it may still overlap the excitation field caused by stimulus *a* in the plane represented.

We may see our way to an explanation of the phenomenon if we assume that the synapses transmitting the impulses which produce the conditioned response to *a* are those having low overcrowding limens. (*Cf.* preceding sections.) If we accept this assumption, we may readily perceive why the application of stimulus *b* (which produces no response itself) *simultaneously* with stimulus *a* produces a partial or total inhibition. For, evidently, the weakened excitation due to *b* and the excitation due to *a* may overlap at some point and, combining, present an intensity of excitation which exceeds the threshold of overcrowding of the effectively conducting fibers. Hence a weakened response, or no response at all, is evoked.

Let us now see whether some light may not be thrown on the

above-mentioned assumption that, as a result of extinguishing *b*, only those paths with low overcrowding limens are left free to conduct. The probable cause for this state of affairs is not far to seek.

The extinction of *b* serves not only to raise the θ values to a degree higher than that existing prior to the conditioning process, since low frequencies are more "harmful" than no conduction at all. It serves concomitantly to raise the synaptic overcrowding limens. (*Cf.* preceding section.) Hence a great many paths and combinations of paths in the vicinity of the efferent projection center are made unavailable to the modifying influences of the excitation due to stimulus *a*, because of their inability to conduct. But not all paths and combinations of paths will be so affected. It is conceivable that of the many thousands present, a certain number, because of some unique quality in their relations to the rest, will escape. These are much more likely to be those paths having low overcrowding limens because these are the ones which are last to be affected by excitations of ordinary intensities, since they are the last to be excited.

Of course, not all the paths and combinations of paths having low overcrowding limens will enjoy this immunity. Undoubtedly, there will be many whose synaptic polarization will be increased, along with that of the other junctions, as a result of the extinguishing process. These should be most numerous in the region between the locus origin of excitation b' and the efferent projection center. There will be some also in the region between the locus origin of excitation a' and the efferent projection center, but there should be fewer here. For, as the extinguishing process continues, the area of effective excitation due to *b* will constrict, and hence this region will become relatively free of its polarizing influence quite early in the process. This region will continue to be invaded by mild excitations, however, as the unreënforced applications of stimulus *b* continue, and additional combinations of low-frequency-limen paths may be expected to become "knocked out" from time to time; but we may readily believe that there will be some so situated as not to

be attainable by the enfeebled excitation due to *b* when this latter is acting alone.

Several corollaries follow from these considerations:

(1) It may be expected that if, under these conditions, the intensity of stimulus *b* is sharply diminished, a response is much more likely to occur, even though both stimuli are presented simultaneously.

(2) If, under the above conditions, the stimulus *a* is presented alone but with sharply diminished intensity no response is likely to occur. (As contrasted to the conditions when there has been no experimental extinction of *b*.)

(3) Extinction of one of the conditioned responses to one of the conditioned stimuli should bring about a weakening of the response to any other conditioned stimulus which has not been extinguished (Pavlov's secondary extinction). This occurs by virtue of the dimished irritability and lowered overcrowding threshold brought about in the areas of overlapping.

The presentation of the conditioned stimulus *a* simultaneously with the unconditioned stimulus is not likely, in the above set-up to inhibit the unconditioned response because of the very probable fact that some of the effective excitation zones involved in the unconditioned response are subcortical.

It might further be expected to follow that if a great many conditioned excitations have been extinguished in the above manner, the absolute amount of cerebral area and mass suffering from this polarization influence would be enormously increased, thus making for a widespread condition of increased synaptic resistance and a decreased limen of overcrowding. It is interesting to note that the Russians, in working with conditions approximating those suggested here, report a tendency for the experimental animal to fall asleep.

These considerations also lead to the suggestion that if several stimuli are conditioned to the same response and unconditioned reënforcement is supplied to each, a greater response and one of longer duration would be elicited by any one of the single stimuli than by the same stimulus if it alone had been conditioned to the response in question.

Further, if two stimuli, *a* and *b,* are conditioned to the same response, stimulus *a* being conditioned first and negatively (that is, conditioned and then thoroughly extinguished), and stimulus *b* later and positively, a greater number of trials would be required for the conditioning of the second stimulus than under ordinary conditions.

It is very appropriate that we consider in this connection the phenomena of *positive* and *negative induction,* as they are termed by Pavlov.

The set-up in positive induction is essentially the same as that given in conditioned inhibition. Positive induction is, in fact, a phenomenon which arises in consequence of certain chronometric relationships. If we have stimulus *a*—as before—positively conditioned to a given response, and stimulus *b* negatively conditioned to the same response, then, it has been shown that, if the application of the inhibitory stimulus *b* precedes the application of the stimulus *a* by some thirty seconds, the intensity of the response to stimulus *a* will be distinctly greater than it ordinarily is, *i.e.,* when the application of the conditioned inhibitory stimulus *b* has not occurred.

This phenomenon is to be considered in the light of the discussion which we have advanced in the preceding paragraphs. When the conditioned inhibitory stimulus *b* is applied, we do not have, as has been indicated already, a complete absence of excitation in the area of overlapping in the vicinity of the projection center. There is some excitation, but insufficient to break through and cause a response. There is also, in consequence, the after-excitation which endures and is subject to the process of temporal decay. Hence, when the stimulus *a* is applied not too shortly thereafter, there will be a reënforcement effect, and, provided the consequent excitation does not exceed the critical threshold of overcrowding, the intensity of the response will be increased.

Clearly, we have to do here with a function in which the dependent variable is represented by the intensity of the response to *a* and the independent variable is the temporal interval between the application of stimulus *a* and stimulus *b*. When this interval is zero, we get no response (conditioned inhibition); as the

interval increases, there should be an increase in response, until a maximum is reached; then there will be a steady decrease, which has as its limit the usual strength of the response given by *a* in isolation.

From this we may readily expect that if the intensity of stimulus *a* were diminished within a limit, the flexion of the curve would come sooner, and its ordinate value might not be zero at any point.

The application of the above principles to the phenomenon of *negative induction* is not obvious at first glance. The phenomenon in question may be described as follows: Stimulus *a* is positively conditioned to the response, as usual; stimulus *b* is negatively conditioned to the same response. If we now proceed to apply, alternately, stimulus *a* and stimulus *b*, reënforcing each in its turn by giving the unconditioned stimulus, we do not succeed in reinstating the response to stimulus *b*. This reinstatement can be accomplished only by refraining from applying stimulus *a* altogether. As soon as stimulus *a* is applied and reënforced, any progress which has been made in reinstating the response to *b* disappears.

If we look again at Figure 5 and consider the various fields, it will be seen that one of the regions of lowered irritability is to the right of the efferent projection center. This region presents, as usual, not only a lowered synaptic irritability but also a lowered threshold of overcrowding. This region is also the region of overlapping between excitation *a'* and the weakened excitation due to *b(b')*. We may express the situation in other words by saying that the cortical zone of diminished synaptic irritability and lowered overcrowding threshold with which we are concerned here is the zone which constitutes the *edge of the furthermost advance of the reënforced* excitation *a'*. This edge, it will be noticed, falls in the field of the excitation due to *b*.

Now, we have already pointed out that, by virtue of the fact that the intensity of an excitation diminishes as it advances from its source, the *outermost* edge of the area of excitation, in consequence of the nature of the function expressed by equation (8) suffers from large increases in polarization. In other words, the

cortical loci which are at the edges of a field of excitation suffer most from increased values of θ and hence from an increased synaptic resistance and lowered overcrowding threshold. As long as the edges of the excitation continue to invade the region in question, the irritability is lowered at the rate indicated by our fundamental equation, and, having attained a limit, is maintained there by subsequent repetition.

We may therefore see that, as long as the stimulus *a* is given and reënforced, this edge of diminished synaptic irritability and lowered overcrowding threshold will continue to exist in the zone to be traversed by excitation *b′* if it is to reach the projection center. This state of increased synaptic resistance and lowered overcrowding threshold in the zone in question has been further abetted by the fact that excitation *b′* has not been reënforced by the unconditioned stimulus.

If, after we have attained this state of affairs, we proceed alternately to apply stimuli *a* and *b,* reënforcing each in its turn by means of the unconditioned stimulus, it is clear that the excitation due to *b,* when reënforced by the excitation due to the unconditioned stimulus, will bring about a frequency value which exceeds the overcrowding limen at the "edge" we have been discussing. This prevents the combined excitations from affecting the efferent projection center; and since the full value of the combined excitations is not realized, there can be no diminution in synaptic resistance in consequence. Hence, there can be no reinstating of the conditioned response to *b.**

This situation, for the reasons indicated, occurs only when *a* is being applied.

* It is reasonable to expect that the phenomenon of negative induction is not readily realized. For, if our considerations are correct, the phenomenon will not occur except when the overcrowding threshold in certain zones has been sharply lowered, and this will occur when (1) the value of θ at synapses is very large and (2) when the synapses are suffering heavy adaptation. For, although we have not considered the influence of this second factor upon refractory phase, we may well believe, in accordance with many of the facts advanced by nerve physiologists, that any factor tending to diminish neural irritability tends to increase refractory phase (Lillie, 23). Hence the phenomenon of negative induction should be obtained much more readily if the time intervals separating the application of the various stimuli involved are made of small orders of magnitude.

If, however, *a* is no longer being applied, and *b* is applied and reënforced (or not applied at all for a period of time), the overcrowding limen—due to diffusion of positive charges at the synapses—will be lowered to a certain degree and the process is begun whereby the combined excitations due to *b* and the unconditioned stimulus gradually succeed in advancing farther and farther in the direction of the efferent projection center. Thus synaptic irritability in the zone in question is gradually increased, and eventually the conditioned stimulus *b* is enabled to evoke the response.

As long as both stimuli *a* and *b* are alternately applied and reënforced, however, the excitation due to stimulus *a* combined with the unconditioned excitation does not permit the limen of overcrowding at this critical edge to recover. Hence, the reënforced excitation due to *b* and the unconditioned stimulus cannot break through, and consequently the synaptic resistance cannot be diminished.

The negative induction phenomenon might be obtained, however, by an alternate reënforced application of stimulus *b*—which has been extinguished—and an unreënforced application of stimulus *a*. Thus, for a short time, we should have a situation in which an unreënforced conditioned stimulus prevents a negatively conditioned stimulus from being reinstated. This state of affairs might be expected to continue as long as the excitation from the unreënforced stimulus *a* penetrates into the critical overlapping area. Having, however, by the appropriate technique, succeeded in reinstating the response to *b,* if this application of stimulus *a* without reënforcement is made to go on for some time, we should expect that *a* itself might no longer be reinstated as long as the reënforced stimulus *b* is applied. Thus the situation would be reversed.

VIII

The Law of Effect and Maze Learning

We proceed to a consideration of the so-called Law of Effect. If we employ Thorndike's classical definition, this may be expressed somewhat as follows: "Those reactions which lead to a satisfying state of affairs are retained in consequence of repetition; whereas those reactions leading to an unsatisfying state of affairs are dropped out."

This type of definition, when applied to explain the behavior of infrahuman organisms has a rather circular look. What reactions lead to an unpleasant state of affairs in the animal? Those that are dropped. What reactions lead to a pleasant state of affairs? Those that are retained. In the introspecting human, if we disregard the epistemological status of introspection, the definition is more satisfactory. It describes, within certain limits, an empirical fact. There is no doubt that in many cases we avoid the unpleasant and seek the pleasant. If we ask an individual why he has failed to perform a certain activity, we are accustomed to hear the answer: "Because I find that sort of thing disagreeable." And we are, as laymen, satisfied with the explanation. Now, it is a well-known fact that in explaining a phenomenon we are forced to begin with certain postulates behind which we cannot go—at all events, at the moment. This is always true, regardless how refined or rigorous the reasoning. In classical physics, for example, the physicist was not expected to go behind Newton's laws of motion. In short, we must start with something which must be taken for granted.

In view of this fact, the question may well be asked: "If, in any explanation, something is always taken for granted, what point is there in devising for a given phenomenon another explanation which attempts to go behind the postulate used in the first explanation, when the second explanation likewise must start with an equally arbitrary postulate?" The obvious answer is

that, by beginning with a more general postulate, we may be enabled to embrace a wider and more varied field of phenomena than we otherwise might. It is, for one thing, a matter of mental economy.

This, of course, is exactly the situation with regard to the Law of Effect. This law is correct within its limits; it is, within these limits, descriptively true. But the limits are small. It can give only a dubious explanation of the fact that an animal, in traversing a maze in quest of food, comes, with practice, to eliminate blind alleys. Using this principle, it is usually said that, since the animal reaches a cul-de-sac, he cannot go on, and in thus being thwarted, experiences an unsatisfying state of affairs. As has been pointed out in a previous paper (11), this may be quite true, but it is also evident that the animal turns, goes out, and subsequently reaches the food. The same reasoning obviously applies to any intervening segment of the path to the food-box. Since food is not reached in any of the intervening segments, these likewise present an unsatisfying state of affairs, and there is no reason why one segment should be dropped rather than another. This is but one of the several obstacles which lie in the path of the Principle of Effect.

Let us therefore try to consider the possible neural conditions. We choose, as a starting point, the following experimental set-up. We have a T-maze, the bottom of which has been wired to permit of the use of an electric current to shock the animal at any time. The ends of the two arms are provided with a door some ten inches from the end. These doors may be closed instantaneously by the experimenter by means of appropriate mechanical controls. The floor of these two compartments are likewise wired, as is the floor of the starting box. The electrical connections are so arranged as to permit of the separate electrification of the various segments of the maze individually. By electrifying the bottom of the starting box, the runway, and the remainder of the maze with the exception of the left compartment, the animal, with practice, learns to go through the left arm to the end compartment, where he is enclosed for a period of several minutes. The process

is repeated many times, so that the activity may be greatly mechanized.

At a given point in the experiment, let us reverse the situation by electrifying the left compartment, and leaving the right compartment non-electrified. The animal runs, as usual, to the left compartment where he is imprisoned for the specified period of time; whereupon he is removed and placed in the starting box for the next run. If this is repeated several times, it is clear that, in spite of the previous mechanizations, the animal will turn to the right arm and go to the right compartment, and will continue to do so as long as these conditions prevail.

What has happened? We need to begin our considerations for this purpose only at the point in the process when the left turn has been thoroughly mechanized. The animal goes to the left compartment as usual, and the door is closed behind him. The electrical stimulation is continued, however, and diffuses the cerebral mass with a continuous stream of impulses. Originally this stimulation had ceased upon entrance into the compartment. This condition had facilitated the fixation process; the reverse condition—continuation of electrical stimulation—seems to destroy it. We are therefore faced with the apparently paradoxical phenomenon of a sudden *interruption* of a flow of of impulses over a given zone of fibers affecting the facilitation of a subsequent response.

The paradox, however, is more apparent than actual. We have seen that the interruption of a flow of impulses has the effect of facilitating a train of responses; the reverse condition—the continued flow of impulses—inhibits the same train of responses. We must take our cue from the latter phenomenon and be reminded of the processes considered in Section I.

The crucial point in the maze running by the animal is the bifurcation in the maze. When he reaches this point, he must turn either right or left. This bifurcation presents to the animal the appropriate set of visual, auditory, tactual and perhaps olfactory stimuli which force him to turn either to one side or the other. It is well known that the preceding kinaesthetic stimuli also are determiners. The previous history of the organism has

been such that this complexity of stimuli force the animal to turn to the left. When he has done so, and reached the end compartment, the predominant and insistent tactual component of this complex group of stimuli—the electric shock—ceases. At least in the original set-up.

We must now examine the processes at work in our second condition, when the electrical stimuli do not cease. We are to be reminded of the process of after-excitation. The nature and distribution of this after-excitation is determined by the preceding kinaesthetic, auditory, visual and olfactory stimuli which had led to the series of responses in question. The stimuli themselves have been removed—all except the electrical stimuli which continue after the left end compartment has been reached. The net result is that these persistent afferent impulses (due to the electrical excitation) spread and overlap the cerebral zones affected by the previous kinaesthetic, auditory and visual impulses and their after-excitations. The consequence is that these zones of overlapping suffer from a strong synaptic polarization.

When the second run is given the animal under the new experimental conditions, the synapses in the zones in question present a diminished irritability to the continued progress of impulses over the various paths leading to the projection centers and effectors which determine a left turn; and as the process is repeated, the irritability decreases still more, with the final result that when the animal approaches the bifurcation, the original sequence of responses cannot occur, and hence a less impervious set of paths is utilized: the animal turns to the right. In fact, were it not for the electrical stimulation received there, the animal would not move from the starting box.

It should follow from the above considerations that, within certain limits, the more prolonged the electric stimuli received in the left compartment, the sooner will the animal change his sequence of behavior at the bifurcation. Not only that, but the more hesitant will the animal be to leave the starting box—that is, the longer the period of inaction and the stronger the stimuli needed to make him leave. It likewise follows that, if the animal is blinded, and all other factors maintained constant, the longer

the arms of the T-maze—the longer the period of time required to reach the end compartment—the fewer the number of repetitions required before the animal reverses his turn at the bifurcation.

We are now in a position to understand why an organism under the stress of the need, will come, with practice, to " choose " the behavior sequence of least action—the behavior sequence involving the least distance and time. It was shown, in another place (11) that learning to run a maze with less and less errors is a specific instance of the Least Action principle. Thus, while repetition is an indispensable factor in learning, it is not the sole indispensable factor. " Motivation " is even more important in a sense; for it " makes " or " breaks " the sequence of behavior which is being performed. It is " motivation " (the reënforcing—or inhibiting—stream of organic or peripheral impulses) which provides the indispensable factors of fixation or extinction.

For the purpose of further elucidation, we shall consider an experiment on white rats previously reported by the writer. The animals were blinded by complete removal of the eyeballs under deep anaesthesia, and were trained in the set-up indicated in Figure 6.

The two arms were of equal length, and the rats were run in this pattern a great number of times, until the habit was thoroughly mechanized. Very early in the training process it was observed that the animals attempted to turn the corner at the elbow shortly before they actually reached it. This is a very familiar phenomenon and constitutes an integral element in the dynamic structure of the learning process. When the habit had been thoroughly mechanized, the wall *a* was removed. All of the animals began by cutting the corner, shortening their paths with each successive trial until they moved in a diagonal path from the starting box to the food box.

In considering this phenomenon, it is evident that the cortical mass is bombarded by a more or less regular train of impulses which, through practice and subsequent lowering of synaptic resistances, eventuate, by way of the appropriate brain zones, into a series of responses terminated at the food box. Each of

these trains of impulses leaves a corresponding after-excitation in the neural mass which is influenced and modified by the excitation of the succeeding afferent impulses. The net result of a run under the experimental conditions we have described is an intricate pattern of synaptic modification in the cerebral zones which have been involved. It is as if a small rivulet of water

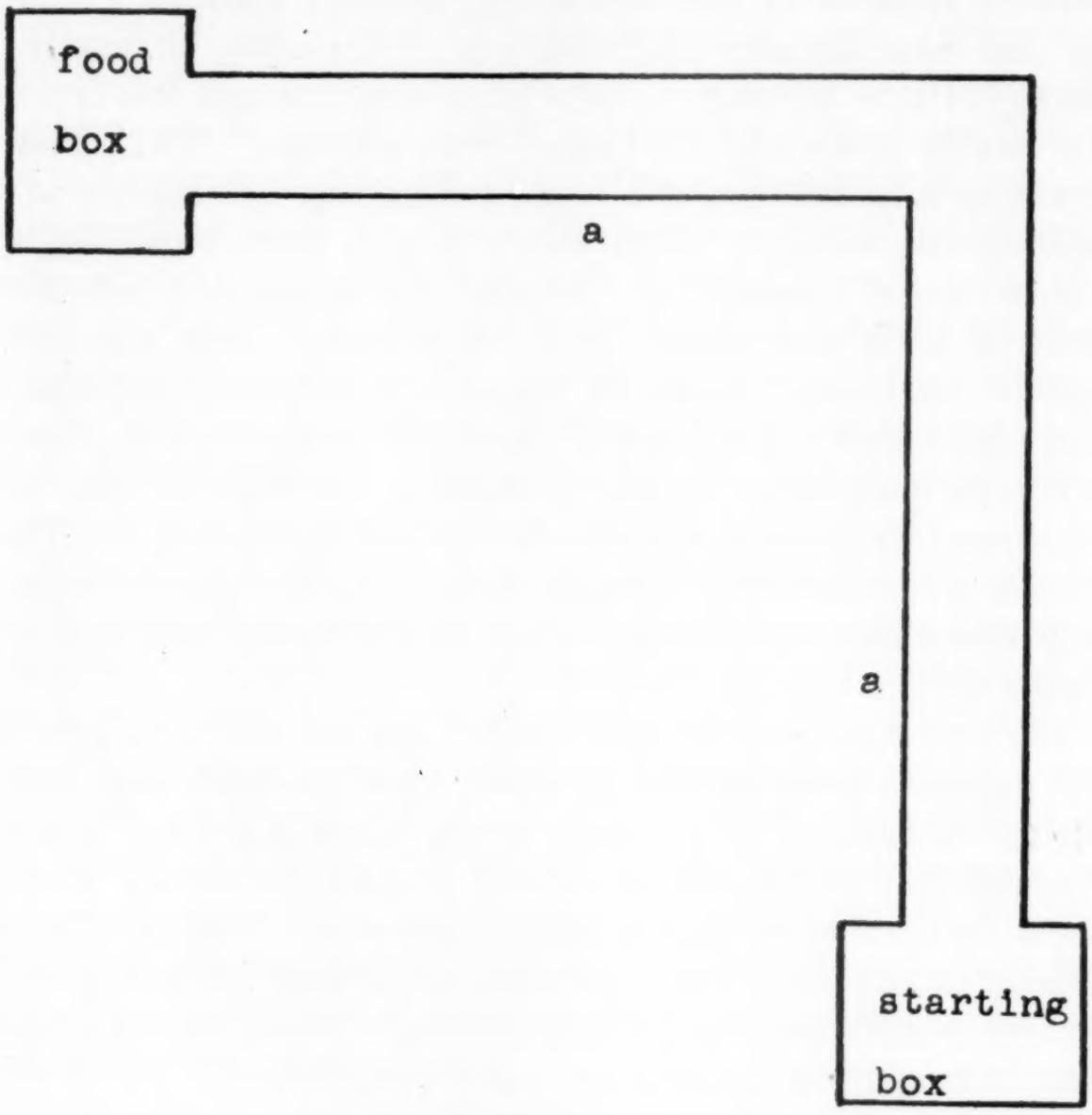

FIG. 6

were turned loose upon a small plot of freshly plowed, sloping ground, and, after flowing for a short time turned off. A number of small rivulets would form therein, mutually influencing each other, until the volume of released water had found its equilibrium. Once this has occurred, the ground presents a modified topography which serves to influence the distribution of a subsequent volume of water.

Leaving analogies aside, however, one of the responses contained in the original training series in the experiment we are describing is the left turn. It is clear from the considerations which were undertaken in connection with the discussion of the simple conditioned response that, as a result of the frequent reënforcement by the organic (hunger) stimuli of the excitations and after-excitations caused by the maze stimuli, the stimuli coming earlier and earlier in the temporal sequence of external (maze) stimuli will, as practice goes on, tend to release the response in question. In other words, the response (the left turn) will be released by the component excitations which precede by a greater and greater temporal interval the excitation which originally released the response. We saw this same principle at work before, when, in Section III, in considering the simple conditioned response set-up, the conditioned stimuli were in the nature of a temporal series *a b c d e*. The present situation is essentially the same.

That the turn does not continue, under the new conditions, to be as sharp with each successive trial is reasonable from (1) the fact of the inertia of forward movement and, (2) the selfsame phenomenon of "anticipation" which does not permit a sharp turn to complete itself before the next forward step is taken.

Evidently, repetition alone is not sufficient to "fixate." It must be abetted by a reënforcing process such as a "drive," and even when it is so abetted, repetition does more than just "stamp in" a response. Rather, it changes responses in accordance with certain dynamic laws. This is made possible by the fact that, ordinarily, a stimulus cannot be considered as isolatedly impinging upon the nervous system. It comes, rather, as a component in a setting, and this setting constitutes a complex determiner of which the particular stimulus is but one part. Thus it becomes possible for the organism to "anticipate" or "redintegrate"; the anticipation is evoked by the temporally prior segments of the complex excitation process. This factor of "anticipation" which we have attempted to discuss in Section III when we were considering the simple conditioned response is to be regarded as ubiquitous in the learning process. We thus see that the dynamic

interaction of the two processes of anticipation and minimum nerve polarization eventually results in the fixation of a minimum path, or, at least, a path approaching the minimum.

It is instructive to follow these mutually influencing processes of anticipation and minimum nerve adaptation in another of the experiments which were reported in the aforementioned paper. The animals were blinded and entered, by way of a starting box, into one corner of a square enclosure whose dimensions were approximately 6′ by 6′. The food box was to be reached by entering a six-inch opening in the diagnonally opposite corner of the enclosure. The animals at first wandered about at random, clinging close to the walls, but with each successive trial the amount of excessive behavior became less, until finally the animals proceeded diagonally across the enclosure to the opening leading to the food box.

If we attempt to envisage the neurological processes occurring with reptition of the runs, it is clear that the first run is one of the most complex from the point of view of the patterns of neural excitation. The traces left in the recebral mass are of a diffused nature, owing to the intricate interplay of diffusion, reënforcement, and inhibition brought about by the " random " behavior and the ever-present organic excitations—hunger. In the second run, we see already the beginnings of the anticipation phenomenon: the animal dashes out more eagerly from the starting box; he does not linger in the starting corner. In other words, there is " a place to go." In neurological terms, those cortical zones which were involved in the response immediately preceding the finding of food have the advantage of the maximum decrement in synaptic resistance. The animal behaves accordingly: he *runs forward,* just as he did when he reached food. The excitation potential, in spreading over the cerebral mass, moves toward the " sink," that is, the region enjoying the maximum irritability, and discharges.

Unfortunately for the animal, purely physical conditions make it impossible for him to stop the steady flow of organic impulses by this mere rushing forward. The goal is distant and located in an intricate coördinate system. Therefore, his nervous system

is subject, in the meantime, to all the contingent influences which prey upon a delicate coördinator. Hence he will wander. If the wanderings on this trial are greater than those on the first trial, they will suffer increasingly from synaptic polarization, and on the next trial will be at a disadvantage. If they chance to be shorter and briefer, the reverse will take place.

In addition to this neural discrimination against the longer of the possible paths, we have, as has been indicated, the tendency to anticipate or "short cut." Thus, if in a given preceding trial there is a "wobble" in the path taken by the animal, it will be straightened out in the succeeding runs, and kept straight by the economy of synaptic polarization. Thus it is that a straight line, under the conditions which we are describing, is approximately achieved. We, therefore, see that the principle of Least Action has as its physiological basis the fact of anticipation and minimal nerve polarization.

These dynamic neural processes which are conceived to underlie the principle of Least Action are the very same ones which we must bring to bear in a consideration of the phenomenon of maze learning.

Let us take, as a starting point, the very simple maze shown in the accompanying diagram, Figure 7.

The animal, motivated by hunger, is placed in the starting box. If we assume the animal to be quite tame and non-diffident, he will presently wander out of the starting box into the runway and walk down toward the bifurcation *o*. As he proceeds forward, the cerebral hemispheres are being continually bombarded by organic, kinaesthetic, visual, cutaneous, and auditory stimuli. If we believe that the neural zones affected by these divers stimulations overlap to some degree, we may readily believe that these various excitations mutually influence and modify each other and produce a unique resultant which is continually changing with time. That the dynamic excitation resultant should change with time follows from the fact that the afferent impulses (particularly from the visual modality) are continually changing and from the fact brought out by our fundamental equations in Section I.

Presently the animal has reached the bifurcation. He cannot go further forward. He must either (1) retract his steps to the starting box, *i.e.*, make a turn of 180 degrees, (2) turn to the right 90 degrees, or (3) turn to the left 90 degrees. The food, of course, lies in the left arm. But let us assume that the animal turns to the right. Neurologically, in terms of our theory, this

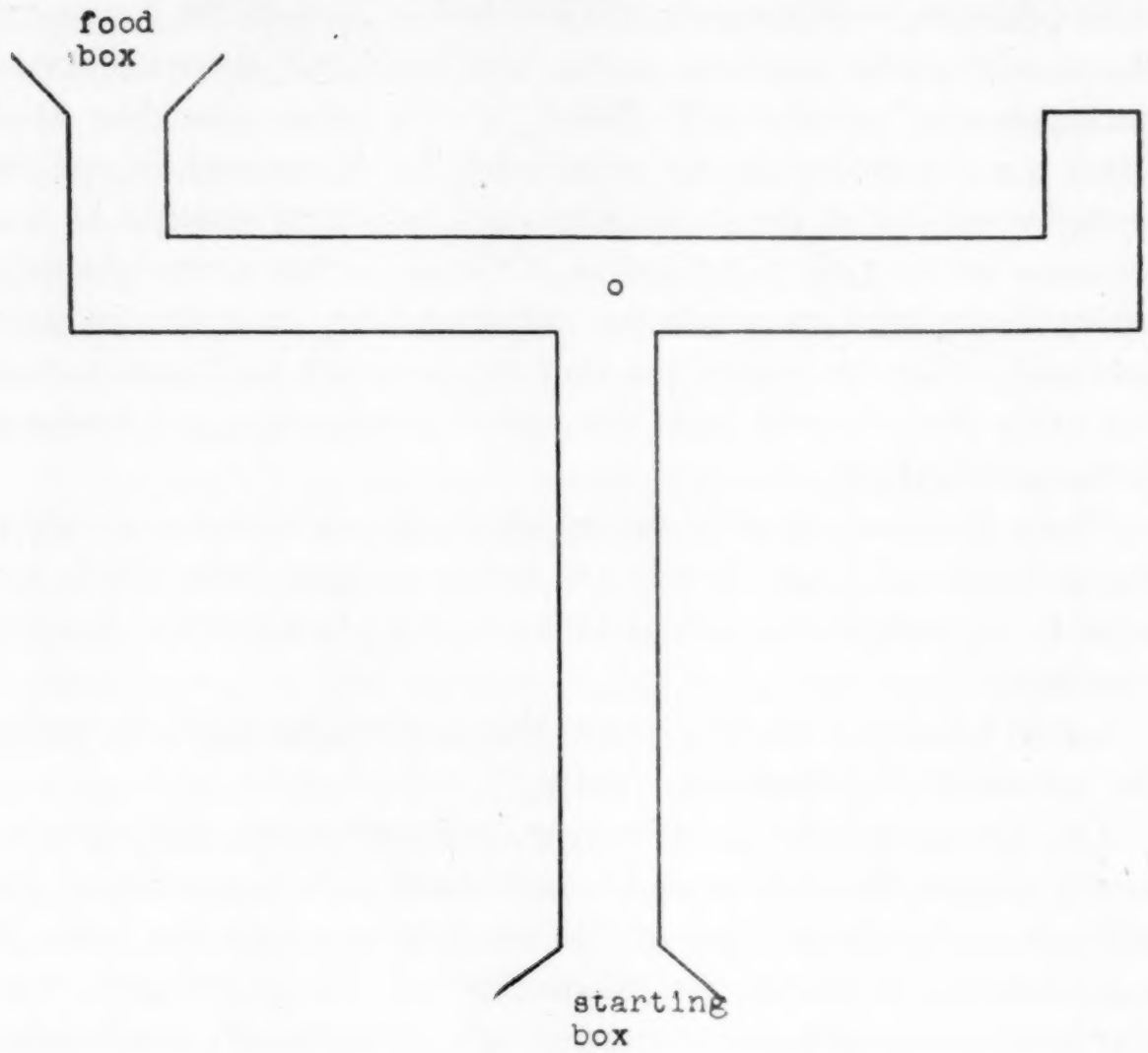

FIG. 7

must mean that the particular distribution and pattern of excitation potential has discharged over nerve zones leading to the efferent projection center governing right turns. Our fundamental modifiability equation would lead us to believe that this heavy discharge of impulses over the zones in question results in a diminution of synaptic resistance within the paths and combination of paths which lead to the efferent projection center. The discharge is particularly heavy because of the reënforcing effect of the organic stimulation (hunger). The result of this dis-

charge would be to facilitate a repetition of its occurrence under the same or highly similar conditions of cerebral excitation.

However, this increase in permeability in the zones in question is somewhat neutralized by the fact that the animal does not reach food. He advances forward in this arm of the maze, and finding, when his head advances beyond the elbow, a blank wall, he eventually turns back. In the meantime the bombardment by organic stimuli of the cerebral zones which have participated in the animal's right turn at the bifurcation continues. For, as we have repeatedly stated, the condition of cortical excitation does not disappear with the disappearance of the peripheral stimuli which caused it. Therefore, there is a reënforcement of the after-excitation due to the exteroceptive and proprioceptive stimuli presenting themselves at the bifurcation by the excitation due to the persisting organic (hunger) stimuli, since, presumably, the zones harboring these various excitations overlap.

The result of this is an increase in the value of t in the modifiability equation, which as we have seen, beyond a certain point increases θ at the synapses involved.

Let us turn to the overt behavior of the animal. He has turned back at the end of the cul-de-sac, and proceeds, let us say, directly to the left arm of the maze, where he continues forward until the food is reached and eaten. The food is of sufficient amount to eliminate the gastric contractions. He is now removed from the food box and returned to the living cage.

Later in the day, when the hunger pangs return, he is placed once more in the starting box. There is reinstated in the cerebral hemispheres the condition of cortical excitation (caused by hunger, handling by experimenter, visual, tactual, auditory and olfactory stimuli afforded by the starting box) which led to the animal's moving forward into the runway on the previous run and eventually reaching the food. However, in accordance with whether, on the previous run, the animal reached food quickly or after a relatively long period of time, the animal will hesitate in the starting box less or more. For, if in the previous run, a long period of time intervened between his starting forward in the runway and the removal of the hunger drive (*i.e.*, the elimi-

nation of the cerebral bombardment by the organic impulses) we must believe that the cerebral zones traversed by the nerve discharge which led, on the preceding run, to the animal's entering the runway, have suffered from a greater amount of neural polarization than they would have if the intervening period had been short.

On this second run, the animal eventually proceeds down the runway and reaches the bifurcation once more. Again he is faced with the alternative of going either to the right or to the left. If, on the previous run, the cerebral zones governing under the conditions of the situation have suffered from a sufficiently large amount of synaptic polarization, the cerebral potential will now discharge over those zones which determine a left turn of 90 degrees. (We ignore, for purposes of simplicity, the possibility of the animal's retracing his steps to the food box. Our situation is already needlessly complex for purposes of unambiguous exposition.) Let us suppose that such is indeed the case, and that the animal makes the turn to the left and, proceeding forward, presently reaches the food box, where the hunger drive is relieved.

Now, a consideration of the relative lengths of the paths taken on the two runs clearly reveals that, for reasons indicated in the preceding paragraphs, a smaller amount of neural polarization will affect the cortical zones involved in making a left turn at *o*. Since the value of *t* is smaller for a left turn, the value of f_s will be much larger in those cortical zones through which a neural discharge leads to a left turn than in those zones through which a discharge leads to a right arm. Hence it is clear that, with succeeding runs, this disparity of excitability between these two zones will be continually increased, with the consequent effect that the right arm will be entered into less and less frequently, being finally avoided altogether. As long, however, as this disparity is not sufficiently great, any contingent cerebral condition may cause the balance to change temporarily in favor of a right turn at point *o*.

Although the maze situation described above is one of the simplest imaginable, the application of the same principles can

be made to mazes of greater complexity, without the introduction of any assumptions other than those we have made during the course of this paper. We shall have occasion to say more about this when we consider the temporal maze problem of Hunter.

Returning to the aforementioned maze problem, let us suppose that the animal by reason of some original disparity in the excitability of the two zones in question persisted in making a right turn at *o* and going to the end of the blind alley before proceeding over the true path to the food. If this behavior persisted for several runs and showed signs of being fixated to a certain degree, we have reasons to believe, on the basis of our principles, that eventually the animal would begin to turn around in the blind alley sooner and sooner, and at length would avoid it altogether. We may believe this from the fact of "anticipation" which we discussed in this as well as in an earlier section. For the animal, upon reaching the cul-de-sac, is forced to *turn*. With repetition, the cerebral excitation state preceding the state which conditioned the turning in the earliest trial would come to be increasingly successful in causing the turn. The mechanism underlying this phenomenon, which we have already discussed at some length in the section on conditioning, is essentially as follows:

The after-excitation from immediately preceding stimuli overlaps the excitation caused by the external situation which leads unconditionally to the response in question (turning) and brings about a reënforcement effect which serves to increase the synaptic permeability of the cerebral zones intervening between the zone bearing the after-excitation in question and the locus of the efferent projection center which governs the turning. With repetition (involving sufficiently large f_s values) the synaptic irritability of the intervening zones becomes sufficiently great to permit the excitation caused by the immediately preceding stimuli to break through to the efferent projection center and bring about the response under discussion. Thus the reaction occurs at a point earlier in time; that is to say, it is anticipated. This process of anticipation or short-cutting continues until a point is reached where, to continue it further, brings it in conflict with the

economy of nerve polarization, that is, brings an increase in synaptic polarization.

We thus perceive that in maze learning, as in other situations, the animal comes to avoid excess distance because of the combined effects of anticipation and nerve adaptation.

The objection is often made by critics of the Least Action hypothesis that the path of least distance is not always the one fixated: that animals very often insist on retaining a bit of "superfluous" behavior in their progress toward the goal. From the foregoing discussions, it is clear why this may occur at times. Evidently, it will occur when the conditions in the cerebral zones governing a "least action" and "excess action" response, respectively, are such as not to present sufficient disparity in irritability in the "correct" direction. If, for example, at a bifurcation, an animal persists in "putting in" a short excursion to the left arm of the maze before going into the "correct" right arm, it is because, as he approaches the bifurcation, the cortical zones controlling the left turn possess—because of constitutional and other factors—greater synaptic irritability than the zones controlling the right turn. This, in spite of the influence exerted by preceding runs. However, the proper discrepancy in irritability may be established by forcing the animal to remain in the "wrong" alley for a period of time before he is permitted to continue to the goal.

IX

Orientation and the Order of Elimination of Blind Alleys in the Maze

This general factor of anticipation which we have been discussing leads to a rather interesting deduction which has been tested recently in an experiment conceived quite independently of any of the theoretical considerations advanced in this paper (4).

If we construct a maze of the following type (Figure 8) where the last turn before food is obtained is to the right, we have reason to believe, from purely theoretical considerations, that the animals will show during the course of the learning process a tendency to make more right turns than left turns. If, on the other hand, the last turn made before food was obtained were to the left, the animal might be expected to make more left turns.

We may persuade ourselves of the plausibility of this deduction by considering that the neural zones involved in making right turns at x_1, x_2, x_3, and x_4 (see figure) do not suffer from as high a degree of synaptic polarization as the neural zones involved in making left turns at these points. For, when the animal has reached the point indicated at x_4 and has made a right turn, he presently reaches the food, and the organic drive is released. Neurologically, this means that the duration of the excitations traversing the neural zones responsible for the right turn has become quite short, and hence a relatively small adaptation effect is produced. This results in a relatively greater permeability (see fundamental modifiability equation) of the neural zones in question. It is true, of course, that prior to making a right turn at x_4, the animal has made previous *right* turns which did not succeed in bringing a prompt termination to the organic bombardment from the stomach and that this has served to increase the adaptation factor in the appropriate neural zones. But it must be remembered that the animal has made probably as many

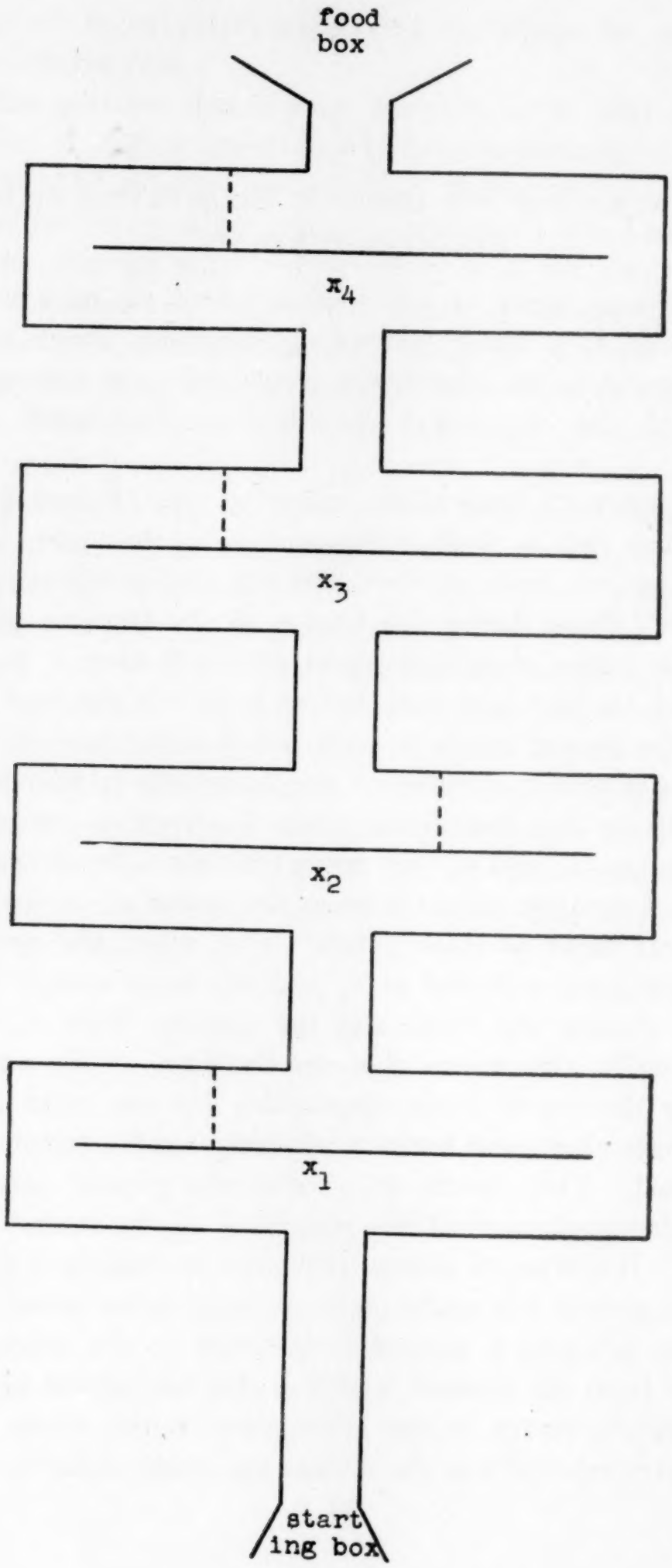

FIG. 8

left turns which likewise failed to lead to prompt termination of the hunger drive. Whether right turns or left turns become fixated at points x_1, x_2, x_3, and x_4, is ultimately determined by the *relative* amounts of synaptic adaptation and polarization involved in making turns of the one kind or the other.

Consequently, when the animal, on subsequent runs, finds itself in highly similar situations as x_1, x_2, x_3, he is likely to go right rather than left. The more homogeneous and similar are x_1, x_2, x_3 and x_4, the greater the tendency of the animal to show this proclivity. Likewise, the smaller the absolute dimensions of the maze, the greater will this proclivity be, because, owing to the shorter durations involved, the greater will be the similarity of the neural conditions involved at x_1, x_2, x_3, and x_4.

Clearly, the preceding considerations are intimately related to the problem of "spatial orientation" in animals which has recently come to demand so much attention. We may say that the above discussion throws considerable light upon the frequently attested fact that animals, throughout the course of their run prior to reaching the food compartment, display tendencies to move in the *direction in which the food is to be found.* And this in the absence of any apparent differential external cues. The capacity of animals for manifesting this extraordinary proclivity on the basis of proprioceptive stimulations can be readily understood in the light of the preceding discussion. For it is evident that the animals are "anticipating" certain crucial turns which lead to the food compartment in obedience to the dynamics of the neural processes which we have already discussed. It is thus readily understandable that maze animals should tend to enter predominantly in those blind alleys which "point" in the direction of the food compartment.

We trust we have made sufficiently clear the point that the central conditions of neural excitation which lead to a definite response on the part of the animal when he is at the bifurcation point in a maze, for example, are not determined exclusively by the exteroceptive and proprioceptive cues which are present for the animal at that point. The neural excitations aroused by exteroceptive and proprioceptive as well as organic (drive) stimuli are modified by the presence of after-excitations of tem-

porally prior stimuli in a manner analogous to that in which the cerebral excitation caused by the unconditioned stimulus is modified by the after-excitation of the temporally prior conditioned stimulus. (See *supra.*) It is the unique dynamic resultant of these various excitation and after-excitation factors which leads to a given response at a particular maze locus. If any of these after-excitation factors are changed, be it by prior stimulus conditions or by the learning process itself, then the response cannot be the same, even though the external stimuli are the same.

Thus, on the basis of our considerations, maze running cannot be conceived as occurring as a series of responses to a series of isolated stimuli. Rather, as a consequence of our conceiving the cerebral cortex as a highly integrated fabric rather than as a network of isolated pathways, and of our insisting upon (1) the presence and persistence of after-excitations to stimuli and (2) the factor of excitation overlapping, we must conclude that the excitation pattern in the cerebral zones caused by a given stimulus *a* is as much determined by the nature of the after-excitation of a prior stimulus as by the nature of the present stimulus *a* itself.

Maze learning, therefore, is a process whereby a series of differential cerebral excitation patterns come to discharge over different projection centers in accordance with (1) the determining condition of minimal synaptic adaptation and polarization and (2) the neural mechanism which, for the sake of brevity and the lack of a better name, we have called anticipation. Because of these ubiquitous neural mechanisms, provided sufficient motivation is given (sufficiently large value of f_s), repetition makes it inevitable that behavior patterns involving the least time and distance should become fixated.

It is clear that these considerations have immediate bearing on the problem of the order of elimination of blind alleys in the maze. Since correct choices of alleys near the food box are the most quickly rewarded by food, that is to say by the elimination of the hunger drive, it is evident that those neural zones mediating such responses are the ones which suffer very little from synaptic adaptation and polarization (since *t* is very small); and the nearer the correct alley in question is to the food box, the less

the value assumed by these two variables. Thus the disparity in synaptic irritability of cerebral zones governing "correct" and "incorrect" responses in the vicinity of reward is distinctly greater and proportionately more in favor of those governing the "correct" responses. Hence the true alleys in the vicinity of the food box are fixated most quickly.

On the other hand, those "correct" alleys which are quite removed from the food box necessarily must be fixated more slowly by virtue of the fact that the cerebral zones which mediate entrance into them suffer from a much larger degree of synaptic polarization. Entrance into blinds at these segments in the maze also brings a great deal of synaptic polarization to the appropriate cortical zones, of course, but if we bear in mind the nature of the relationships expressed by our modifiability equation it becomes clear that there is greater discrepancy between durations of—let us say—three seconds and five seconds than between five seconds and ten seconds. Hence a much greater number of repetitions are required to bring about sufficient disparity in the synaptic irritability of cortical zones mediating entrances in true and blind alleys situated in the early segments of the maze.

These polarization and adaptation factors are complicated, however, by at least two others. These latter two may counteract the influence of the former to the extent of preventing the backward elimination of blind-alleys from appearing as a clean-cut experimental phenomenon. It is these two factors which have tended to militate against any experimental unanimity in regard to this phenomenon.

One of these is the factor of anticipation which we have just discussed in connection with the problem shown in Figure 8. It is evident that for certain types of mazes, this factor may assume such importance as to seriously mangle the unambiguous appearance of any eliminating tendency. Therefore, future experiments purporting to examine this phenomenon must needs be carried out with greater attention to the geometry of the maze situation.

The second factor which offers a not inconsiderable influence is that of proactive and retroactive inhibition, and this we shall discuss in a later section.

X

Acceleration in a Straight Runway

The preceding discussions lead us to believe that if a straight runway were involved in reaching food, the animal would tend, provided the factor of motivation remained constant, to run faster and faster in the runway on successive runs. We may expect this if we make the plausible assumption that increased permeability or lowered synaptic resistance is associated with a more vigorous and a more rapid response. Hence, the rate of speed may be expected to increase from trial to trial as it approaches some limit. If we assume that there is a direct linear relationship between irritability at the synapses and rapidity of the response, we may reasonably expect that the increase in velocity from one trial to the next for any given segment of the runway will be negatively accelerated or sigmoid. (*Cf.* Section I.) The actual slope of this function will vary for different portions of the runway, being sharper for those segments which are nearer the food box, and flatter for those nearer the starting box. This latter inference may be expected from the relationships expressed in equation (12).

It may be expected further that, shortly after the practice period has begun, the velocity of the animal on any given run will accelerate as the food box is approached, when it will decelerate as the food box is almost reached. This acceleration is to be expected from (1) purely physical factors of momentum and (2) the fact that the neural processes which occur in answer to stimuli immediately preceding the removal of the internal drive have suffered less from synaptic polarization. (See equations and curves in Section I.)*

* It is interesting to record here that the above discussion, as well as all that follows, was written prior to the appearance of Professor Hull's (18) brilliant articles on this subject. Thus, the same conclusion has been reached independently, and from theoretical considerations in both cases.

If the sides of the runway were painted with a series of discernible, non-repeated patterns, the acceleration would be more pronounced. In other words, the less homogeneous the appearance of the path, the greater the acceleration.

It is evident further that if we were to plot a curve of the relationship between the velocity of the animal and his distance from the food box for each successive run, the slope of the curve will become less and less sharp with each successive run. For, with each successive run, the absolute velocity with which the animal traverses the pathway will increase, hence the range between the extreme time values (t_x–t_o) will become less. Furthermore, the above considerations are influenced by the purely physical factors. When the animal is approaching his limit of velocity in running the pathway (*i.e.,* in the very late stages of the training process), we may expect a very sharp acceleration at the very beginning of the run, and the remaining distance will be run at approximately constant speed.

It is also suggested, from the phenomenon of anticipation which we have already discussed, that with continuation of practice, the point of deceleration in speed on the part of the animal will begin slightly sooner as it approaches the food box, within some limit determined by the economy of nerve polarization.

We have thus far considered the above situation as involving animals with normal exteroceptive receptors. The problem becomes somewhat different when we consider a blind, deaf, anosmic and cutaneously anesthetized animal. By eliminating the exteroceptive organs, the heterogeneity of the situation is removed, hence the excitations which occur are exclusively organic and kinaesthetic. The kinaesthetic excitations themselves achieve a maximum of homogeneity in the above situation where a straight runway is used. Now, although it should be granted that with motivated and rewarded practice the animal's speed will increase gradually with increasing repetitions, the problem arises as to whether, during the course of any given run, the animal will show acceleration or deceleration as he approaches the food box.

Since the source of proprioceptive stimulation is practically

homogeneous because of the simple type of behavior involved—mere running forward—and the source of organic stimulation may, to all intents and purposes, be considered as likewise homogeneous, we may expect increasing synaptic polarization to occur in the cerebral areas involved as the animal proceeds forward to the food box. (See Figure 1 in Section I.) The polarization effect will not be confined to the cortical zones, but will occur also in the proprioceptive receptors themselves. In view of the preceding considerations bearing upon animals with normal exteroceptive receptors, we may, therefore, expect in this case that deceleration will occur as the animal moves forward toward the food compartment. There will, of course, be a sharp acceleration at the beginning of the run as the animal leaves the starting box, and this will continue until the animal has attained the maximum momentum. Once this is reached, however, the animal slows up more and more as the food compartment is approached. Clearly, this deduction holds only for straight runways of sufficient length. If the runway is too short, the polarization and adaptation factor is masked by the purely physical aspect of the animal's momentum.

XI

Latent Learning

At this point it would be pertinent to indicate that the tenor of the above discussion would lead us to expect the results reported by Kuo in his now classical experiment (22). It will be remembered that Kuo's rats were quick to eliminate the route leading to a compartment wherein they were imprisoned for a considerable period of time before they were permitted to proceed to the food box. The interim of imprisonment in the food box gave ample opportunity for the continuous and persistent stream of organic impulses in the hungry animal to effect considerable degree of nerve polarization.

It is interesting to consider in this connection Blodgett's experiment on " latent " learning. During the days of delayed feeding, the animals showed very little improvement in learning to run the maze. As soon as food was introduced in the food box, however, the curve took a sharp drop. It is significant to note that in this experiment (3), the drop in the curve subsequent to the introduction of food in the maze was much sharper than the corresponding segment of the curve for the control group of animals. As Blodgett remarks, there was some evidence of " latent " learning.

In considering this phenomenon of latent learning, it should be indicated that, prior to the termination of the period of delayed feeding, the curve for the experimental animals did drop, although very slowly. This indicates that there was a slight increase in synaptic irritability despite the long duration of the organic bombardment. Or, to be more precise, the slight diminution in number of errors is to be looked upon as a *combination* of the increased polarization of synaptic junctions mediating diffuse reactions (fear reactions, " startle," etc., etc.) and decreased polarization of synaptic junctions mediating entrances into true paths. This, of course, is always the case.

We have not mentioned the former component heretofore in the interest of simplicity. This differential process is hypothecated upon a constitutional difference in the frequency of the nerve impulses mediating these diffuse "emotional" and "tension" responses, and those mediating simple locomotion through the maze.

Now, the crux of our problem here is to account for the fact that on a given run—the run succeeding that when food is first placed in the maze—the drop in the learning curve for the experimental animals (the delayed feeding group) was much greater than for the control animals. If we term the run on which food was immediately found, the nth run, it is clear that this run served to increase the discrepancy in the synaptic polarization between those regions mediating "true" entrances and those mediating "false" entrances to an extent not achieved by any of the preceding runs. Nevertheless, if we follow Blodgett, we are to construe the result of the experiment as indicating that the runs preceding the nth did have an effect which was not manifest until after the animal had been fed at once in the maze on the nth run.

Now, while this phenomenon of latency seems very mysterious, in reality it is not more so than the "latency" which occurs in connection with the establishment of the simple conditioned response. For, evidently, the critical process of synaptic *de*polarization goes on for some time before there is any overt manifestation of it. During this period, the process may be considered as sub-threshold with reference to the response involved. Thus, the conditioned and unconditioned stimuli must be applied, in some cases, several, in others, many, times before there is any overt evidence of conditioning. Yet these early applications with their "sub-threshold" effect are an indispensable part of the process. Once this threshold is surpassed, the conditioning process, as it is manifested by changes in the size of the conditioned response, goes on apace. It is likewise to be believed that for "unsophisticated" animals, the very first few trials in a reasonably complex maze serve merely (1) to increase polarization at those synapses mediating the diffuse reactions and tensions mentioned previously and (2) to bring the depolarization process

at the "relevant" synapses somewhere near the threshold level. Once this stage is reached, the error curve drops very rapidly. Inspection of most error curves for complex mazes and for naïve animals shows this to be the case. Curves obtained under these conditions do not drop sharply at the very beginning—as we would have classes in elementary psychology believe.

In the case of Blodgett's experimental groups of animals we have an instance of this important "subthreshold" process enduring over a period of many trials, owing to the effect of the delayed feeding. (Large values of *t: cf.* equations in Section I.) We may conceive that the duration of the organic bombardment (the animals were fed an hour after the run was completed) was such that the net result of the non-rewarded runs did not succeed in doing much more than bringing the depolarization process to the critical threshold. If this be true, we should expect that increasing the number of non-rewarded runs would not materially increase the amount of drop in the error curve occurring between the first and second rewarded runs. Inspection of Blodgett's data reveals this to be the case. For the amount of drop in errors for the group of animals having seven unrewarded runs was not reliably greater than the amount of drop for the animals having but three unrewarded runs.

In the light of the above considerations, it becomes evident that the drop in errors on the trial succeeding the first rewarded trial should be unusually sharp.

It would seem that we are justifiably led to expect from the general tenor of the preceding discussions that if a group of animals were run through a maze for a number of times in a non-hungry state, without a reward, and then were placed in the same maze when in a hungry state and rewarded, the drop in the curve subsequent to that point would be less sharp than for a comparable group of animals which had been run under precisely the same conditions, with the exception that they be in a very hungry state. We are inclined to this deduction by virtue of our fundamental modifiability equation. (Greater value of f_s in the second group of animals.)

That this deduction may not be very far from the truth is

indicated by a recent paper of Tolman and Honzik (29), wherein the group of hungry and non-rewarded animals gave a distinctly superior form of learning curve than a group of less hungry and non-rewarded animals. This is quite in accordance with our discussion.

We may, further, expect some form of negatively accelerated function to obtain between the number of repetitions required to master a given maze and the amount of temporal delay in feeding. This function should hold only within fairly restricted limits of temporal delay, because beyond a certain point repolarization does not increase. (*Cf.* Section I.)

It is interesting to consider at this point what should reasonably be expected if, in the midst of the learning process—let us say, the mastery of the maze—the motive were suddenly changed. For example, if the animal has made considerable progress in the mastery of a problem with the aid of the sex drive, at a given point in the process, he is placed in the experimental situation in the absence of the original drive but under the influence of a previously unemployed drive, say hunger.

It is clear that under these conditions the locus of the organic component in the stimulus mass has been changed, with a consequent change in the locus and intensity of the excitation zones in the cerebral mass. This should work to the disadvantage of the behavior process in question, although the disturbance should not be very great, in view of the fact that the excitation zones in the cortical mass for the various afferent impulses probably overlap to a greater or less extent.

This, in fact, was found to be the case in Elliott's investigation (6) where, at a given point in the learning process the hunger drive was substituted for the thirst drive. On the day of the substitution, there was found a reliable increase in the amount of errors and time. This sudden rise in the curve, as we might expect, quickly disappeared under the new conditions. It is evident that the amount of this increase must vary with the degree of difference in the loci and spread of the two excitation fields, or, to express this differently, with the degree of overlapping of the two excitations.

DeCamp's and Yoshioka's results (5, 32) on the preference

for the shorter alleys leading to food on the part of white rats is also readily deducible from the above considerations. The longer path involves a greater delay in the satisfaction of the drive, hence in the removal of the continuous internal component of the stimulus complex. The greater synaptic polarization of the corresponding cerebral zones mediating the behavior involved in traversing the longer path reacts therefore to its disadvantage, with the result that the shorter path will be chosen, provided the conditions of the difference limen are satisfied. We should also expect that if a hungry animal, in obtaining food, is faced with the alternative of being imprisoned in either of two compartments for varying periods of time, the animal will eventually show preference for the compartment with the shorter imprisonment time, provided, again, the difference in the temporal intervals fulfill the conditions of the difference limen.

The function which governs the difference limen in this type of set-up is deducible from the principles which we have been discussing. Since the curve of repolarization beyond the reversal point, is in general, negatively accelerated (see Fig. 1), the difference in length between two absolutely shorter paths needed to establish a consistent discrimination is less than the difference needed for two absolutely longer paths. That is to say, if the difference between two repolarization processes must have a value y before the behavior involved in one is consistently preferred to the behavior involved in the other, clearly, the absolute difference between the two distances will be smaller if the distances themselves are of a smaller order than if they are of a larger order.

The above considerations would lead us to believe that if an animal is to master a maze, the task is perceptibly easier when the absolute dimensions of both true and blind alleys are smaller than when they are greater. Thus, if two comparable groups of rats were run in two mazes of identical pattern but varying in their absolute dimensions by the ratio 2 to 1, we should expect the group running in the smaller maze to master the problem more quickly. Further, a maze of a given pattern will probably be mastered more quickly when the blind alleys are long than when they are short.

XII

The Temporal Maze Problem

It is both interesting and instructive to consider whether the neurological considerations which have been advanced in the previous sections are applicable to the temporal maze problem which has been investigated by Hunter (19).

Let us suppose that in the set-up shown in the accompanying diagram (Figure 9), the compartment *o* serves in the capacity of food-box as well as starting-box. The task facing the animal is the following: Leaving the starting-box, he must run down the central alley *c,* turn to the *right* at *x,* return by alley *r,* enter the central alley *c* once more, turn *left* at *x,* return by alley *l* and then enter the runway to the food-box wherein he obtains food. Hence the animal is faced with the relatively simple sequence *rl* at point *x*.

The process to be adopted in training the animal to solve this problem is of some importance. Therefore in attempting a problem of this kind, it would be eminently desirable to choose a method of training which presented the minimum difficulties for the animal subject. We shall not, however, concern ourselves here with a discussion of the relative merits of the various procedures which might be adopted, although the theoretical principles which we have posited would lead us to pronounce certain techniques as distinctly more satisfactory than others. We will assume, rather, that, regardless of procedure, the animal has solved the problem, and concern ourselves with a discussion of the possible neural processes which could render such a solution possible.

There are two points in the maze which are critical for the animal:—*x* and *f*. Of these two, *x* is perhaps the more critical because it offers the less differential cues. Whereas the subject approaches *f* from two different directions on the two occasions, he approaches *x* from the same direction and therefore the external stimulating conditions are the same in the two cases.

Our problem may, therefore, be considered solved if we can discover what differential internal conditions serve to cause the

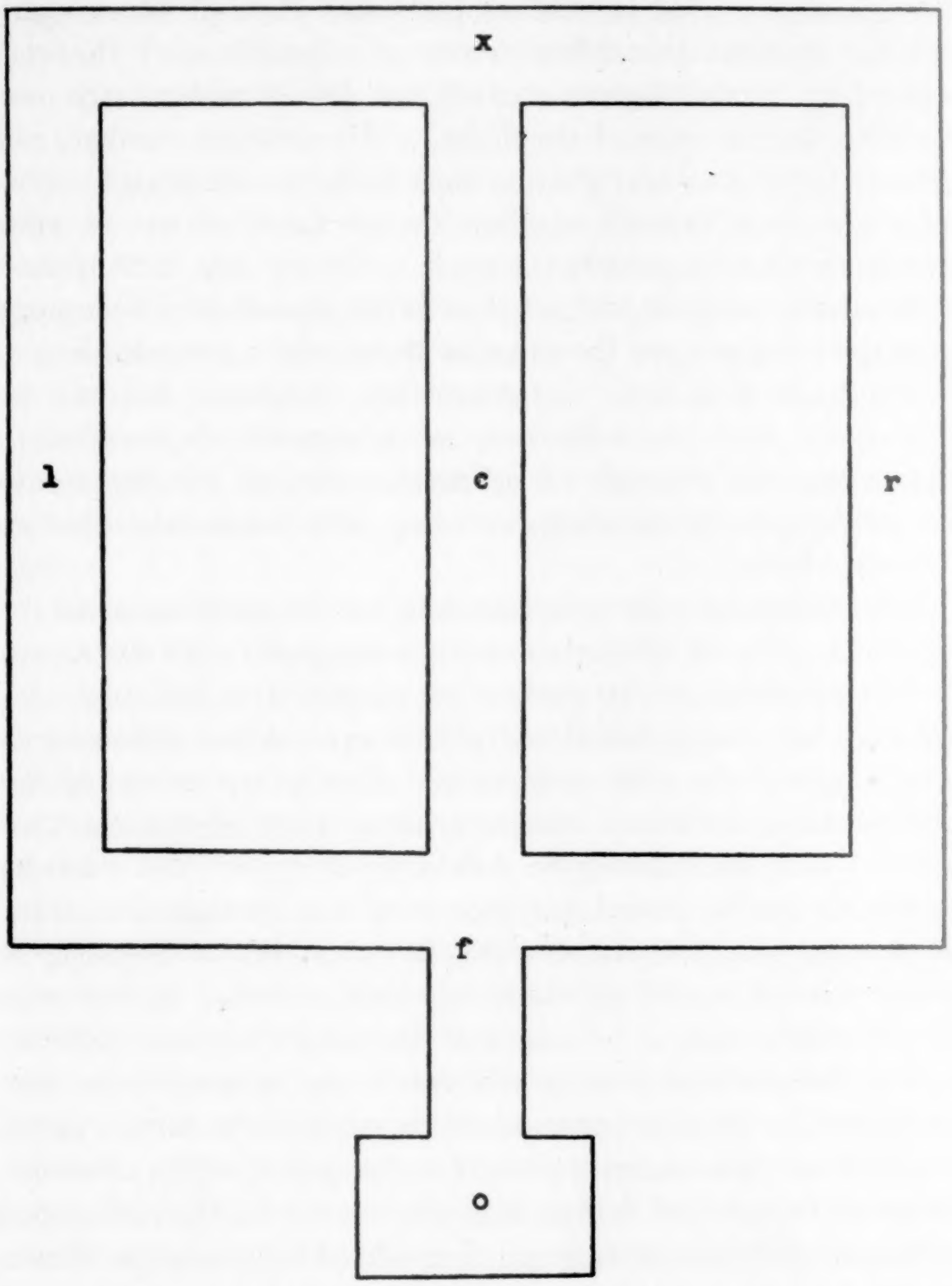

FIG. 9

animal to respond differentially to the same external situation on the two occasions.

Insight is achieved in regard to the matter, if we bear in mind the rôle of after-excitation in learning situations of all kinds.

Let us follow the animal's behavior as he sets out on a given successful run. He leaves the starting-box and runs forward through alley *c* until he reaches point *x*. Here he turns *right,* and the appropriate cerebral zones are flooded with the corresponding proprioceptive excitations (to consider only one modality, for the sake of simplicity). He proceeds further, and again is forced to turn right; he must make two more right turns before he finds himself in alley *c* once more. Now he runs forward until he approaches *x* again. During this latter period of time, the external and proprioceptive stimuli which impinge upon the receptors are the same as those which preceded immediately his turning *right* on the previous occasion. But now he turns left. Since it is the duty of a scientific explanation to acknowledge the principle of sufficient cause, we are faced with the question of determining how the same cause may produce different effects.

The answer, of course, is that it is *not* the same cause in the two cases. But an adequate answer is not given until the source of this difference and its method of operation is indicated.

As has been suggested already, the source of this difference is to be sought in the after-excitations. Clearly, the nature of the cerebral states on the two occasions during a run when the animal makes a turn at *x* are quite different. For, the first time he approaches *x,* the animal has proceeded in a straight line from the starting-box; the second time, however, he has come by a path involving several right turns which preceded his entrance into the central alley *c*. As we have shown already in a previous section, the cerebral state which determines a reaction in any given point in the maze is not a sole function of the stimuli which present themselves to the organism at that point; rather, the conditions of the cerebral field at that moment are further influenced and modified by the already existing after-excitations for which temporally prior stimuli are responsible. It is not surprising, therefore, that the animal should reach differently in his two approaches to the point *x*. The contemporaneous exteroceptive, organic, and proprioceptive stimuli involved at point *x* may be the same for the two occasions, but the cortical fields are not.

And it is the cortical fields which immediately determine learned reactions, rather than afferent impulses. Rigorously considered, therefore, in light of these factors, it is no more surprising that the animal should make successively different turns at point *x*, than that he should make a right turn to one shade of grey and a left turn to a distinctly (for him) darker shade of grey.

Since the differential cortical states in the two approaches to point *x* are mediated by differential after-excitation states, it must follow that if these cortical after-excitation states are eliminated or reduced, the animal should encounter greater difficulty in mastering the correct sequence of responses. This effect can be brought about experimentally. For, since the intensity and efficacy of after-excitations diminishes with time, an increase in the temporal interval separating the animal's first and second approaches to the critical point *x* will serve to bring about this effect.

This may be achieved quite readily by increasing the absolute dimensions of all the alleys in the maze. If the alleys are sufficiently increased in length, the cortical after-excitations left by the preceding right turns will be so diminished by the time the animal reaches point *x* the second time that their use as differential conditions will be appreciably diminished. Hence the animal is more likely to be thrown into confusion.

These same considerations would lead us to believe that if, in an experiment wherein two strictly comparable groups of animals were forced each to learn the above problem, the one group forced to master the problem in a maze of larger absolute dimensions would require more repetitions for complete mastery than the group faced with the same pattern of smaller dimensions.

Greater difficulties are faced by the organism when, with the same pattern shown in the above diagram, the sequence of turns at *x* becomes *rrll*. The chief obstacle, our theoretical discussion would lead us to believe, would show itself on the first *l* turn. Let us follow the possible neurological processes involved, assuming that the animal has solved the problem.

The animal emerges from the starting box and runs down the central alley *c* to point *x;* the cortical state consequent upon

this run forces the animal to turn right. (The reader will bear in mind that we are not discussing here the process whereby this is accomplished. We must regard it simply as a *fait accompli,* trusting that an adequate account of the learning process has been given in the previous sections.) The animal now enters alley *r,* makes the appropriate right turns and finds himself in alley *c* again, where he proceeds down to *x*. This run through *r* has worked its effect upon the animal's cortical mass and produced a dynamic state of excitation which is a resultant, as we have seen, of the after-excitations produced by the various previous turns as well as the excitation engendered by the contemporaneous stimuli at point *x*. The nature of the learning process has been such that the animal once more turns to the right at this point. In fact, the problem demands that he repeat precisely the cycle of behavior and proceed once more to point *x*. Now he must turn to the *left!*

Clearly, the resources of the animal's cortical system are now being sorely pressed. For whereas, in the preceding simple alternation problem, the cortical states present when the animal found himself at *x* were distinctly different on the two occasions and thus provided adequate basis for a differential response, the cortical states which now present themselves on the second and third occasions, respectively, when the animal finds himself at point *x* would seem to be very similar if not almost the same. They cannot be very different for the simple reason that the behavior which precedes the animal's third choice at point *x* is no different from the behavior which preceded the animal's second choice at *x*. What mechanism would possibly be at the disposal of the nervous system to enable it to successfully negotiate this difficulty?

The first point to be kept in mind is the fact that the field of cortical excitation cannot be quite the same at the time of the animal's second and third approaches to point *x*. Although the two immediately succeeding cycles of behavior are the same, the cerebral states cannot be because of the factor of the process of synaptic polarization. In other words, the distribution and intensity of the cerebral excitation field when the animal

approaches *x* the third time will be somewhat different from its distribution and intensity on his second approach. In the former case, because of the factor of polarization it will be less widespread and intense. In short, the impulse frequencies will be smaller.

We may now be reminded of the considerations which were advanced in the section on the "trace" conditioned response. If the animal is to successfully negotiate the proper sequence of turns at this point, we must believe, on the basis of our principles, that the process of learning has been such as to give one set of nerve frequencies access to one efferent projection center (right turn) and the other (weaker) set of nerve frequencies access to another projection center (left turn). The neural conditions are highly analogous to a situation in differential conditioning, where the animal gives one conditioned reaction to a stimulus of a given intensity, and another conditioned reaction to the same stimulus of a weaker (or stronger) intensity. Whether or not, therefore, the animal can overcome this difficult point in the sequence of turns ultimately depends on whether the structure of his particular nervous system and the nature of the external problem is such as to provide adequately differentiated patterns and intensities of cortical excitations in the second and third approaches to point *x*.

Once the animal has successfully turned left on his third approach to *x*, his subsequent path (a series of left turns) provides sufficiently differentiated cortical excitation to enable him to turn *left* once more when he approaches *x* for the fourth and last time. Having done this, his entrance into *f* rather than *c* is guaranteed by conditions analogous to those obtaining when he turned into *l* at *x* for the first time.

It may be clearly seen from the preceding considerations that the process of learning to react successfully at points *x* and *f* presents great difficulties for the animal. We have already noted that the only difference in the cerebral excitations leading to differential turns at point *x* on the animal's second and third approaches, respectively, at that point is based on the factor of synaptic polarization. The difference between the two excita-

tion states, therefore, is not very large, and we have here again the phenomenon of a smaller nerve frequency producing a response which a higher nerve frequency fails to produce. The mechanism making possible such a condition we have already discussed in connection with our consideration of the trace conditioned response.

On the other hand, if the animal were forced, in the above maze situation, to learn the sequence of turns *rlrl,* he would encounter less difficulty. A glance at the sequence of paths which this problem entails and a brief consideration of the corresponding neural excitations involved makes it obvious that this should be so. Clearly, each approach at point *x,* for example, is accompanied by adequately differentiating cerebral states. The first turn at that point to the right, is preceded by the straight run through the central alley *c;* the second turn at that point, to the left, is preceded by the run in *r* entailing a series of right turns; the third turn, to the right again, is preceded by the run in *l* with its series of left turns. (Thus we have two fairly distinct cerebral states leading to the same response, viz., the right turn at *x.* These two states are the resultant of the straight run from the starting-box and the run in alley *l,* respectively. This presents no particular difficulty to the organism, however; it is, in fact, analogous to the process of attaching the same conditioned reaction to two different conditioned stimuli.) The fourth turn at point *x* is once more to the left, and is preceded, as it was on the second occasion of the animal's approach to *x,* by a run in alley *r.* It is true, of course, that the cerebral state mediating the animal's first left turn at *x* is not identical with that mediating his second left turn at that point, even though the objective path pursued by him prior to making these two turns is the same. Nevertheless, it is evident that it is much simpler for two highly similar cerebral excitation states to lead to the same response than it is for them to lead to diametrically opposite responses.

In fact, our theoretical considerations would lead us to believe that in the single alternation problem which we have considered, the more difficult point for the animal is at *f* rather than at *x.* For at *f* the animal must make differential responses which are

preceded by the running of the same objective path (viz., after his second and fourth turn at *x*). The situation involved here, therefore, is analogous to the situation facing the animal when he must make his first left turn at *x* on the double alternation problem *rrll*.

Our discussion, therefore, based upon the theoretical considerations which we have advanced throughout this paper would lead us to conclude:

(1) That the sequence *rlrl* in the temporal maze is much more readily mastered by an animal than the sequence *rrll*.

(2) That the sequence *rrll* may be mastered by an animal provided it possesses a sufficiently complex cortical mass. That is to say, provided the fibers in the cortical mass have, among them, a sufficiently great range of overcrowding limens.

For, since in the sequence *rrll*, the behavior preceding the second *r* turn and the first *l* turn is the same and therefore the pattern of excitation in the cortical mass is *qualitatively* the same, the only source of differentiation must lie in such differences as are effected by the polarization factor. This, as we have seen, consists in a diminution of the frequencies in the neural impulse trains. Thus the animal must make an *r* response under the cerebral condition of a given impulse frequency and an *l* response under the cerebral condition of a lesser impulse frequency, while the patterns of cerebral excitation remain qualitatively almost the same. This may occur successfully only when the fibers mediating the *l* response possess an overcrowding limen sufficiently low to prevent them from being activated by the (greater) frequency which determines the *r* response. (See discussion of the conditioned "trace" reaction.)

XIII

Retroactive Inhibition and the Spaced vs. Unspaced Methods of Learning

The phenomenon of retroactive inhibition and the general superiority of distributed as compared to massed practice may be readily considered in the light of the discussions in the preceding sections.

If we are reminded of the fundamental relationship which is indicated in our fundamental equations, we see that the greater the duration of application of a given stimulation, beyond a certain point, the greater the degree of synaptic polarization and adaptation, or, in other words, the less the degree of increased synaptic irritability achieved.

Now, the phenomenon of retroactive inhibition is usually described as the fact that the interpolation of activity between the acquisition of a given activity and its later recall results to the detriment of the latter. In other words, the amount forgotten during a relatively "empty" period is distinctly less than the amount forgotten during a "full" period of the same duration.

If we examine the factors in these two situations, it is clear that a period full of interpolated activity, provided the region of excitation due to this latter partly overlaps the region due to the original activity, makes for a greater amount of synaptic repolarization in those cerebral zones which have been and, by virtue of the after-excitation, still are, functioning as a result of the original activity. The combined effects of the "motivating" stimuli and the external stimuli constituting the interpolated activity itself, result in an increase in the duration variable in the equation, and hence in an increase in the amount of repolarization.

Although the occurrence of the second or interpolated activity with its accompanying motivation factor tends to increase in

certain zones the intensity of the after-excitation consequent upon the first activity, the effect of the prolonged duration may eventually more than compensate for an increase in irritability of the synapses in the overlapping area. That this is highly probable may be seen when we consider that the factor *t* is steadily becoming larger, whereas the value of the intensity of the nerve processes in that part of the zone mediating the original activity where overlapping is occurring is not correspondingly increasing.

If, however, the interpolated activity is instituted an appreciable interval of time after the original activity has ceased, the intensity of the after-excitation due to the latter has greatly diminished or vanished; hence the degree of possible reënforcement is less, the amount of overlapping area involved becomes less, the distribution of synaptic repolarization is curtailed, and consequently the recall of the original activity is affected to a less degree.

It follows further from our discussion that the amount of retroactive inhibition varies as some direct function of the duration (within limits) of the interpolated activity. It also varies as the degree of "similarity" between the motives involved in the original and interpolated activity, as well as the degree of "similarity" of the activities themselves, within certain limits.

We have introduced, in the preceding sentence, the phrase "within certain limits." For it is obvious that if the interpolated activity is identical in nature and motive with the original activity, the performance of it (or rather the continued performance of it) later does not make for inhibition, but facilitation. In such a case, the cerebral conditions during the performance of the interpolated activity approximate quite closely the conditions which preceded it, and the intensity of excitation becomes sufficiently great to offset to some degree the duration factor. As the two activities become increasingly "dissimilar," however, there will be a tendency to less and less overlapping, which will make for intensities which will be overbalanced to a greater and greater degree by the duration factor. At some point in the "similarity" scale, the amount of overlapping becomes sufficiently small to permit of a flexion in the process, and beyond this point the

amount of retroactive inhibition continually decreases until it approaches zero, when there is no overlapping. It is interesting to compare these considerations with those of Robinson (26) and with the results of an experiment by the present writer (12).

The enhanced value of t and an insufficient intensity of excitation in the overlapping zone, therefore, are the most important factors in the phenomenon of retroactive inhibition, since the advent of an interpolated activity imposes a continuous bombardment of impulses of relatively low frequency upon the overlapping area; this makes for a distinct increase in the value of the repolarization at the synapses; hence the area involved is in a state of partial "block" when the original activity is reinstated, and diminished efficiency results.

It is suggestive to note here certain relationships which would seem to exist between the phenomenon which we have been discussing in the preceding paragraphs and the experimental findings reported by Lashley in his celebrated extirpation experiments. Lashley indicates that the loss of function in a particular activity acquired by the animals through training varies directly as the amount of extirpated cortical material. Our discussion above leads us to believe that the amount of retroactive inhibition varies as some direct function of the size of the area suffering from synaptic polarization as a result of overlapping of the zones of cortical excitation. Evidently, in retroactive inhibition the blocking effect is in a sense temporary, since it is relieved to a degree by the process of synaptic adaptation and polarization recovery—a state of affairs quite different from that obtaining when the cortex is mutilated. However, the two phenomena seem to present certain fundamental features in common, and these characteristics are to be linked with what was discussed in Section II, where we considered the relationship between the strength and duration of the response to a given stimulus and the amount of available cortical mass.

The preceding considerations anent the possible processes underlying retroactive inhibition lead us to believe that retroactive inhibition is but one phase of a mutual inhibition process: the other phase to be termed proactive inhibition. For, if retro-

active inhibition is caused by diminished synaptic irritability induced by the overlapping of two zones of excitation, it is clear that the interpolated activity must suffer as well as the original activity. Thus, it should follow that the acquisition of a function is more difficult when it follows close upon the heels of a preceding activity than when the immediately preceding interval is relatively vacant.

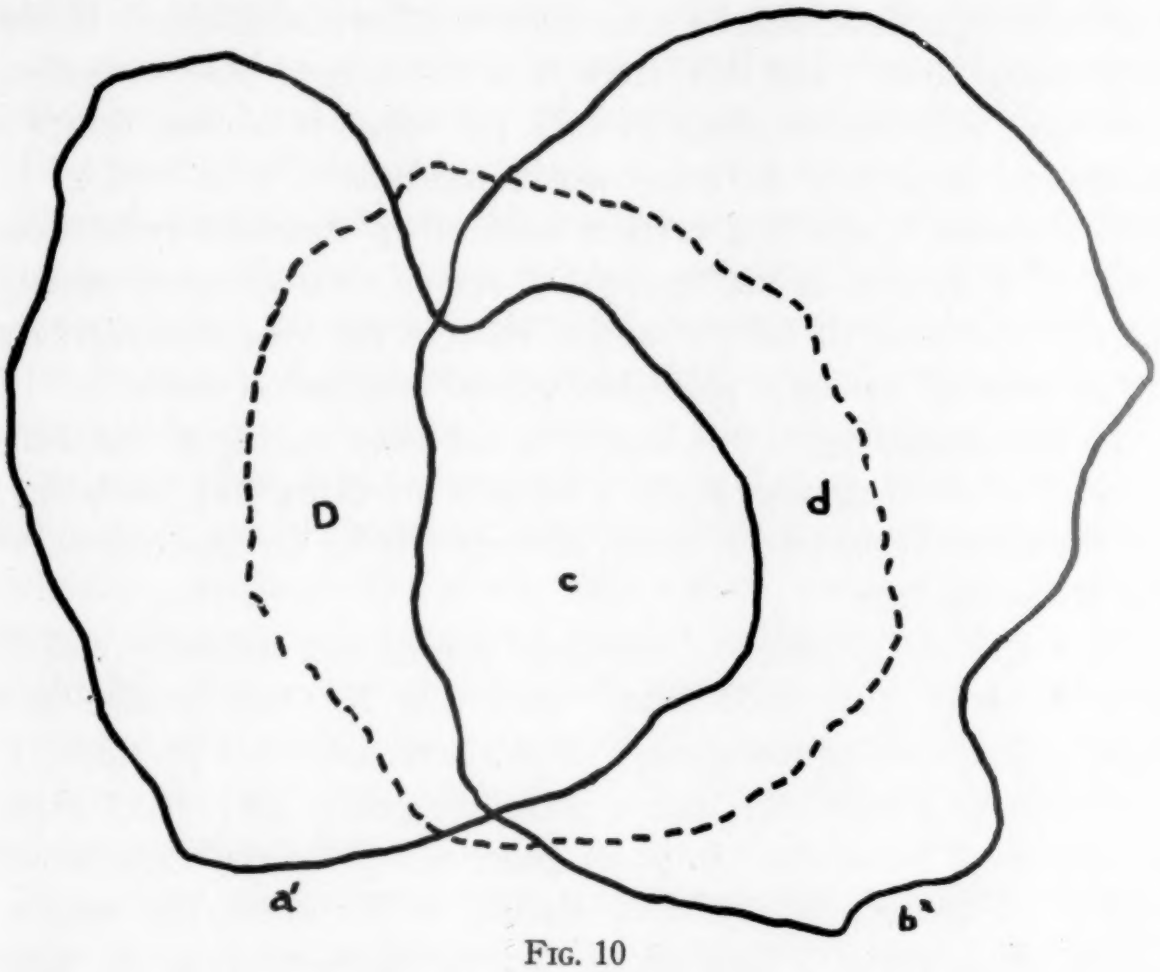

FIG. 10

The question arises at this point as to the relative strength of proactive and retroactive inhibition under comparable conditions. From purely theoretical considerations arising from our preceding discussions, we may make out a case for the greater strength of retroactive inhibition.

This may be illustrated by Figure 10 wherein is shown the result of the overlapping of the after-excitation from one stimulus with the excitation from another immediately succeeding it. a' shows the distribution of the after-excitation from stimulus a after a given period of time has elapsed since its inception and when the stimulus has impinged in isolation upon the organism.

b' is the excitation area aroused by stimulus b when it impinges in isolation upon the organism. c is their common sphere of influence. d represents the area of "secondary diffusion," that is, the result of the overlapping of after-excitation a' and excitation b'.

By studying this figure it is evident that the area b' has more to gain (better, less to lose) by the overlapping than area a'. For, although there will be considerable polarization effect in the overlapping area c and this, since it is common to both activities, interferes with one as much as with the other, it is clear that the process of secondary diffusion works more harm to a' and to b'. For, whereas it effects a certain amount of desirable reënforcement to b' (see d in the figure), it serves to disproportionately increase the value of t for a′ and to enlarge the area, and increase the amount of synaptic polarization and adaptation there.

To experiment upon this matter is not easy, owing to the difficulty of controlling adequately a number of disturbing variables; but the above deduction is susceptible of test by an experiment of the following type:

Two lists of syllables of identical length and difficulty (*sic!*) may be used. Call these List A and List B. We require also three strictly comparable groups of subjects, a control group (C), a retroactive group (R), and a proactive group (P). All three groups must be at their limit of practice before the experiment begins. Exposure of syllables should be by automatic means. Group R is given 5 repetitions—let us say—on List A, then immediately is transferred to List B and given 5 repetitions on it. When this has been done, the group is once more transferred to List A. The amount retained is determined. Group P is given 5 repetitions on List B, is immediately transferred to List A and given 5 repetitions; then waits for a period of time equivalent to that transpiring in the giving of 5 repetitions, and transferred to List A. Group C follows the requisite routine for a control.

An experiment of the above type, in order to corroborate our deduction, should show a superiority of Group P in the amount retained.

It is further evident that the above discussion applies to the

fairly well-established fact of the superiority of the spaced as against the unspaced method of learning. For, in the unspaced method, the value of t in our fundamental equation is at a maximum, and hence the value of the polarization and adaptation factor is very high. Now, in the spaced method the value of synaptic adaptation is diminished by the period of recovery, whereas this cannot take place in the unspaced method because the process of stimulation continues. Hence, under the latter condition, the value of increment in synaptic depolarization becomes smaller, and a greater number of repetitions are needed to reach a given level of efficiency.

The relationship between the serial position of an item in a list of nonsense syllables to be learned and its relative difficulty (assuming that the list is otherwise homogeneous) in recall is also to be seen as fundamentally influenced by the factors we have been considering. Evidently, items at the beginning of the list suffer most from retroactive inhibition and least from proactive inhibition; items at the end of the list suffer most from proactive inhibition (particularly when the spaced method is used); whereas items in the center of the list suffer from both. We should therefore expect that items in nonsense syllable lists which are found at the beginning and end of the list will be the ones which are learned first; while those in the middle of the list will tend to be learned last.

We might also expect that, since proactive inhibition, other things being equal, is less extensive in its effect than retroactive inhibition, those syllables at the end of the list will be learned a little more quickly than those at the beginning. This deduction may be expected with justification when the spaced method of learning is used.

If the unspaced method of learning is used, however, and the list of syllables is read from beginning to end without pause until it is learned, we may expect to find less discrepancy in difficulty between syllables in various parts of the list than we would if the spaced method is employed.

It is relevant at this point to indicate that the above factors of proactive and retroactive inhibition have considerable influence

on the process of maze learning in both animals and men. It will be recalled that in a previous section our theoretical deductions anent the hunger drive led us to conclude that there should be a backward order of elimination of blind-alleys in animal maze learning. It is clear that the process described at that point purporting to account for this phenomenon is complicated by the factors which we have been discussing in the above paragraphs. Therefore, the actual order of elimination of blind-alleys by the animal will be the algebraic resultant of these three factors (if we ignore the influence of the "anticipation" factor already discussed). Consequently, we cannot expect that there be a perfect backward-order elimination of blind alleys by the animal during maze learning.

Now that we have considered these three pertinent factors as they influence the phenomenon in question, it might not be amiss to venture an approximate evaluation of their resultant effects. Let us indicate the influence of the two factors of retroactive and proactive inhibition on the difficulty of the serial order (always assuming that the items are otherwise equal in difficulty) by the curve *a* in Figure 11. Curve *b* in the same figure represents the influence of the drive factor discussed in Section IX when it is acting alone. The resultant of the interaction of these three factors we may represent by the dotted curve *c* in the same figure. This curve, as we have pointed out before, however, is further somewhat complicated by the factor of "anticipation" which is difficult to control. Nevertheless, if we may exclude from our considerations this disturbing factor, our discussion would lead us to deduce an order of blind-alley elimination roughly represented by this curve.

This deduction should be expected only when the spaced method of learning is being employed. If the unspaced method of learning is used, the situation is further complicated (see *supra*). The optimum distribution of practice for investigating the above deduction would probably be one or two runs per day (one in the morning, one in the afternoon in the latter case), the animal being well fed in the maze at the end of each run.

The question may arise at this point as to why the influence of

the drive factor should be so predominant in the maze situation for the animal, and yet not show itself in nonsense syllable lists for humans. For clearly, if we compare curve *a* in Figure 11 with curve *b,* it is evident that the serial order of difficulty in learning nonsense syllable lists is more nearly represented by curve *a.*

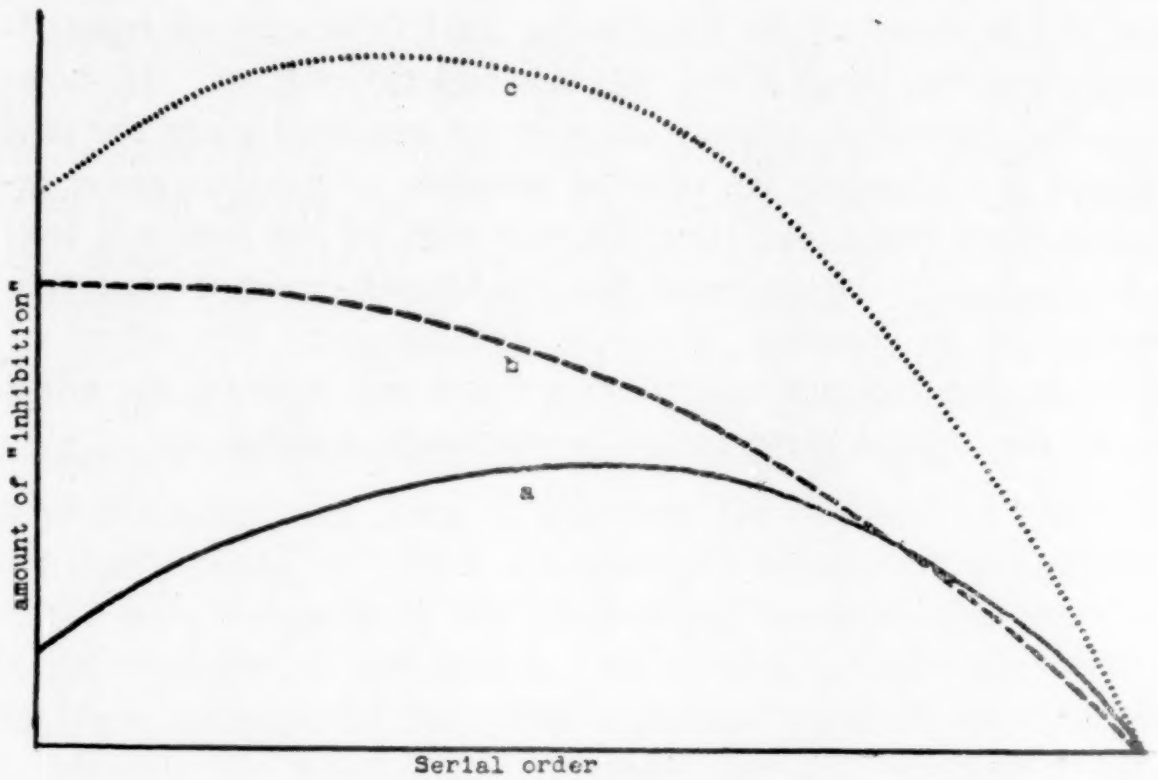

FIG. 11. The curves should not be considered as purporting to represent the absolute amounts of inhibition involved in the various serial positions. We are interested here merely in the *distribution* of the relative amounts of inhibition within the series. Hence the absolute level of the curves have no significance. The curves are predicated upon an optimum use of the spaced method of learning.

b represents the influence of the drive factor; *a,* the influence of the retroactive and proactive factor; *c,* the resultant.

The discrepancy between the serial order of difficulty in nonsense sylable lists and maze blind-alleys is to be found in the different conditions governing the "drive" or motivational element in the two situations.

In learning nonsense syllable lists, the fact that the subject has read through the list from beginning to end does not release him from the presence of the "drive." The drive or "tension" is not released in the subject until the syllables have been successfully repeated without error. Therefore, we may expect the cortical zones of the subject to be continuously bombarded even

after he has read through the list, despite the fact that a spaced method is used. This is not likely to be the case with an animal learning to run a maze. Once the animal has gone through the maze, he enters into the reward compartment and is fed; the drive is relieved, and the cortical bombardment and its concomitant adaptation and polarization effect is removed. Hence those items near the food end of the maze suffer least from cortical repolarization, whereas those at the opposite end suffer most. If, however, the internal tension or drive is not removed when the unit attempt is completed, the relative amounts of synaptic polarization in those zones mediating the two ends of the task will tend to be equalized. Hence, when this condition is realized—approximately—as in learning nonsense syllable lists, the effect of proactive and retroactive inhibition stand out, whereas the effect due to the organic drive factor is practically eliminated.

XIV

The "Pleasure Principle"

The foregoing discussions throw interesting light on the time-honored principle that we "pursue the pleasurable and avoid the displeasurable," with its equally time-honored corollary "the burnt child dreads the fire." The attempts which have often been made to "explain" the phenomena indicated by the above-mentioned aphorisms on the basis of simple substitution of responses due to conditioning can hardly be considered satisfactory. In the case of "the burnt child dreads the fire" the orthodox conditioned reflexers tell us that the sight of the fire unconditionally calls forth in the child the response of reaching and manipulating it. This automatically produces a stimulation of the pain receptors in the hand which produces the unconditioned response of withdrawal. As a result of this association, the mere presence of the visual stimulus (sight of fire) serves to prevent the child from reaching for it. Evidently, this explanation is based upon concepts which are too loose and which therefore serve to obfuscate the nice distinctions which really obtain in the behavior. For, what we should really expect if the conditioned response formula is to be applied rigorously is a sharp withdrawal and avoidance reaction on the part of the child each time he sees the fire. What we usually observe is *a failure* on the part of the child to reach for the fire. This criticism has already been made by Professor Tolman in a recent article. He says:

"It appears that what our conditioned response friends really do is to divide all responses into two sorts—positive and negative. And they argue that in a trial-and-error situation the acts which get learned are those which result in bringing the animal into the presence of further stimuli to which positive responses are already attached. And the acts which do not get learned are those which result in bringing the animal into the presence of further stimuli

to which negative responses are already attached. These resultant positive and negative responses get conditioned back to the cue stimuli. It must be noted, however, that the positive and negative responses which thus get conditioned back may in concrete terms be as different from the original responses from which they are supposed to be derived as entering is from eating or as not-entering is from jumping back and squealing. But this last is a little point which is not stressed by the theory. Our conditioned response friends are truly both serpentish and dove-like." (28, p. 248.)

On the basis of our principles, however, we are in a position to appreciate and accept these behavior differentia and to predict them. We may deduce that "painful" stimuli will lead to elimination of the reactions which originally procured them if we assume that the frequency of nerve impulses actuated by stimulating pain receptors is low. Hoagland (17), in his study with frogs, finds, in fact, that stimuli which elicit "struggling reactions in the intact animal and which in man would be 'painful' give rise to nerve impulses which are of longer duration and low frequency." Adrian (1), furthermore, obtains similar results and adds that nerve discharges from "painful" stimuli are of longer duration than from weak stimuli.

On the other hand, we may deduce that responses giving rise to "pleasurable" stimuli will be maintained and fixated if we assume that "pleasurable" stimuli are those stimuli which cause receptors to actuate impulses of high frequency. Again, referring to Hoagland, we find him declaring that stimuli which are epicritic, such as touch (*e.g.*, with a feather) and light pressure give rise to nerve impulses which are rapid and of short duration.

The bearing of this upon the matter of the instinctive and the acquired in the behavior of organisms is suggestive. Obviously, much more data of the type reported by Hoagland must be accumulated before any reliable insight may be obtained into this ubiquitous problem. But it is interesting to note that many of the so-called instinctive urges, such as sex, must be influenced to an appreciable, if not a large degree, by experience. If the culmination of the sex act were constitutionally accompanied by

intense pain instead of pleasure, it is not certain that the sex drive would be sufficiently strong to insure the continuation of the species.

Similar considerations may be advanced in regard to our penchants and idiosyncrasies. Probably no one would be willing to admit that children come into the world with a native craving for sweets. A person brought up in a culture where no candy or sugar existed would certainly not miss it in the sense that he would miss food. Nevertheless, if such a person were introduced to sweets, very probably he would wish to *renew* the experience and might even develop a "drive" for it. The reverse process might occur in connection with condiments. In a very fundamental sense, both these phenomena would involve a native or instinctive component.

Therefore, in the light of the considerations and relationships which have been advanced in the foregoing pages, there would be a definite neurological basis for the statement that a great many "instinctive tendencies" assert themselves for the organism as a result of experience. And if our fundamental postulates are not in error we should be able to predict what the direction of the tendency will be in a given case.

XV

Insight and Trial and Error

When we come to the matter of "insight" we are faced with a problem which is beclouded by an infelicitous term. The term insight as applied to learning has led many people to think of learning as being essentially dichotomous in nature: thus we have (1) trial and error learning, and (2) "insightful" learning, and these two forms are sometimes considered as diametrically opposed to each other. This dichotomous manner of conceiving the learning process, it would seem, works harm for both members of the dichotomy, in that probably each becomes more obscure in consequence. For while the term "insight" is vividly descriptive, it tends to create difficulties rather than banish them. This fact, as such, is not to be held against any term or concept, but it would seem that, neurologically speaking, the term leaves us stranded high and dry with a bare phenomenon and with no way of relating it to the less dignified "trial and error" form of learning.

We may define insightful behavior as *adequate* behavior occurring in response to a situation which the animal has not faced before; "adequate" is, in turn, defined as the *minimum* or approximately minimum behavior needed under the conditions to relieve some need, as hunger, sex, etc. Thus, if an organism, under the stress of some need or "drive," is placed in a situation which is capable of solution, and the animal at once or very shortly responds with a more or less complicated train of behavior which brings adjustment or relief with a *minimum of "random" or waste motion,* we say that the animal has behaved very "intelligently" or "insightfully."

Now, most laymen and some psychologists would be willing to submit the criterion that in such a case the "animal knows what it is about." Unfortunately, the phrase "knowing what one is about" has no meaning except in the human world, where,

upon analysis, it turns out to refer to the fact that, if the subject is interrupted during his activity and questioned, he will be able to inform the questioner *verbally* as to what he is going to do. Since the animals are not capable of verbal expression, it would be gratuitous to project this meaning into the situation and to define the situation in these terms. For a rat's *first* entrance into a blind alley may be just as "business-like" as the ape's fitting two bamboo poles together. The ape, however, has behaved "insightfully" because he has adopted one of the solutions of minimum effort; the rat behaves "uninsightfully" because he has not.

As we have just seen, insight, by definition, is shown in a situation which the animal has never faced before; or, at all events, in a situation where the *arrangement* of the constituent elements has not been faced before. The animal must "solve" the problem with a minimum of activity, which means, with a minimum of "random" or "trial and error" activity. Now, there is no formal or *a priori* reason for believing that behavior by "insight" should not occur; and that it *does* occur has been shown at least by Tolman and Honzik and Maier since Köhler's classical experiments appeared. Why the earlier exponents of behaviorism should have insisted that all behavior is, genetically considered, of the random movement and chained response type is not at all obvious. Formally, it is no more mysterious that a blow on the patellar tendon should evoke the knee jerk than that a banana hanging from the middle of the ceiling should prompt a hungry ape to pile the available boxes under it before climbing up to reach it.

Certainly, an "insightful" act by an animal high in the phylogenetic scale should offer less of wonderment to the psychologist than the marvelously complex native trains of behavior which are observed in the insect world. One is just as mysterious or as obvious as the other. We are quite willing to accept the phenomenon of the patellar reflex, for example, as evident by virtue of a superficial acquaintance with the phenomenon—an acquaintance which is reinforced by misleading analogies concerning electrical conductors leading to solenoids, relays, and so

on. Yet, as a matter of fact, even the patellar reflex is, strictly speaking, extraordinarily complex in its details when we consider the fact of reciprocal innervation and the orderly and properly timed sequences of impulse volleys which bring about that response. That we should glibly accept the mechanism of reflex response as self-evident and a more complicated and less common reaction as fundamentally different and mysterious is itself a commentary on the psychology of the " obvious " : that is obvious which we have experienced most frequently or which has strong similarity with things or events which we have experienced very frequently.

From these purely formal considerations one may be convinced that the question as to whether learning is fundamentally a matter of trial and error or a matter of insight is just as sensible as the question as to whether water is formed by the same causes as air. A given piece of behavior is uniquely determined by the dynamic and structural conditions obtaining within the organism and the environment. Under a given set of conditions (of which the structure of the nervous system is obviously an important constituent) a particular type and sequence of behavior occurs; it is, in an absolute sense, neither " insightful " nor " random." In fact, the adjectives themselves are a piece of anthropomorphism which the experimenter, from his vantage point of a superior nervous system and sensory equipment, unwittingly commits. In a very real sense, all behavior is insightful: *as insightful as the obtaining conditions permit.* This sentence should not be construed as intending to slur over the *actual* differences in behavior which two animals may show in the same situation when impelled by the same needs. The differences are certainly there, but they should be considered as differences of the same fundamental order.

The problem then resolves itself into a matter of determining under what conditions an organism will behave in a manner which involves the minimum waste of effort and time in satisfying a need when it faces a situation for the first time. Some of these factors may be enumerated: (1) the extent of previous experience in similar situations, (2) the degree of cerebral develop-

ment, (3) the intensity of the internal drive, (4) the amount of sensory equipment, (5) the spatial and temporal relationships of the constituent parts of the external situation to each other, and so on.

It is, of course, unfortunate that we are not enabled to look into the nervous system and study the trend of events as we may study a microscopic field under a cover glass. At the present time, therefore, we are not in a position to give a minute account of the intricate and mutually influencing dynamic processes in the nervous system that lead to a complicated train of behavior events which neatly and economically satisfy the organism's need in a given situation when he meets it for the first time. So that, if the question is asked, "Why does an ape pile boxes one on top of the other in attempting to reach a suspended banana when he has never seen a box before?" we are forced to answer: "For the same reason that he sneezes when there is an irritation of the nasal passages."

The question as to whether these extreme cases in the "insight" or economy scale may occur in the total absence of relevant past experience is a purely factual one and does not concern our purely formal considerations. The question answers itself when we have experimentally investigated and rigorously controlled the various conditions which determine behavior in general. Whatever the result may turn out to be, the fundamental proposition remains that the behavior which has occurred is a unique function of the total conditions which were present. Thus, an ape may fumble with a problem which a man may solve at once; but the ape behaved as "insightfully" as the particular conditions obtaining in his case permitted. The argument of those who see in chained conditioned responses forged by experience the universal paradigm of all behavior is an argument taken from observation of a special class of conditions; it is in consequence too limited and inadequate. So it was with the ancient astronomers who postulated crystal spheres to hold the planets in their orbits because of their experience with the limited conditions obtaining in rotating bodies on the earth's surface.

At the risk of belaboring the point, we wish to repeat that the

factual differences between "random" and "insightful" or economical behavior is not to be slurred. The difference is a very real difference and may be ascertained in any event. We are merely persuaded of the fact that the difference is one of degree rather than one of kind. The distinction between insightful and random behavior has served, and will continue to serve, a very important purpose. If animals, under rigorously controlled conditions, are shown to be able to adapt themselves economically and efficiently to a situation the first time it is met, we have forthwith a fact which leads us to infer that the organism is susceptible to a greater complexity of conditions than had ordinarily been believed. That the organism responds in ways which are richer and more complex than we had thought. In short, our range of study and observation is thereby increased, and the laws which we may propound, concomitantly increase in generality and power. The question, then, is not whether an animal has "insight," but what *degree of "insight" the conditions obtaining in a given experimental situation enable the animal to show.* Or, stated in other words, what the conditions must be in order that the animal show the maximum insight or economy.

The problem may be expressed in a generalized manner by writing it in equational form:

$$n_b = F(d, h, e, s) \qquad (13)$$

where

n_b represents the amount and *direction* of behavior involved in any given attempt to satisfy a need
d represents strength of drive
h represents hereditary endowment of the organism
e represents experience
s represents structural and dynamic complexities of the external situation.

It goes without saying that our function is one involving a number of complex variables, the nature and units of which it is the task of research to determine. Let us, for the sake of simplicity, set the factors

$$(d, h) = u$$

We may therefore write

$$n_b = F(u, e, s)$$

We may precise our expression further by writing it as

$$n_b = F_\theta\left(\frac{s}{ue}\right) \qquad (14)$$

For, evidently the value of n_b varies as some direct function of *s* and as some inverse function of the other two independent variables *u* and *e*.

Now, as has been indicated, although we are not in a position at present to obtain quantitative measurements of the variables involved, for purposes of purely formal consideration, inspection of equation (14) is sufficient to indicate that "insightful" behavior and pure "trial and error" behavior are merely extreme values of the same general function. Thus, when the factors in the right side of the equation attain certain values, n_b becomes a *minimum*. As these same factors assume other values, the dependent variable n_b increases over a continuous scale, and the "trial and error" quality of the behavior becomes increasingly pronounced.

Clearly, the relationship of the independent variables *u*, *e*, and *s* to each other in equation (14) make it possible for n_b to approach a minimum in any of several possible ways. n_b assumes a minimum value when (1) *s* becomes very small, (2) *u* becomes very large, (3) *e* becomes very large, (4) when *e* approaches zero—*provided s is sufficiently small and u sufficiently large.*

This fourth condition evidently describes what is ordinarily meant by insight.

If we single out variable *e* for discussion, clearly n_b decreases, up to some limit, as *e* increases in value. So that, in any given situation facing the organism where the value of the *e* variable is large, the value of the dependent variable n_b will be correspondingly smaller. This is evidently the case when we compare the behavior of an animal threading a maze on the twentieth run with his behavior on his first run.

As we have already indicated, however, the nature of the relationship between the variables permits the value of the dependent

variable n_b to become or approach a minimum when the independent variable *e* approaches zero. This becames possible only when the other independent variables *u* and *s* assume "optimum" values. Nevertheless, it is highly probable that in any given situation the minimum value of *e* which is compatible with a minimum value of n_b never becomes zero.

If we now turn our attention to the *s* variable, it is clear that there must be a limit to the complexity of the external situation to which the animal will adapt itself with the minimum effort upon the first trial (*i.e.*, where *e* approaches zero). As is well known, a great deal of our prided insightful behavior as humans is, as a matter of fact, a form of mental trial and error. This is merely another way of saying that the finite organization of our nervous system is such that it is not susceptible *simultaneously to all* of the intricate factors which the situation may present, and hence portions of it are reacted to separately and consecutively until the organism attains a solution which is then overtly expressed in behavior. Whether all insightful solutions in infrahuman organisms are of this type it is difficult to say. If we may repeat ourselves again, it is, however, conceivable that there is a type of situation which, falling within the limits of a given organism, affects it as a unique dynamic pattern and leads to a unique performance of maximum economy upon the first trial. It is merely a matter of discovering what this maximum value of *s* may be which is compatible with a value of *e* of approximately zero when *u* has a given value. In view of the factor *u* in the equation, it is evident that, for any given organism, the value of *s* which is compatible with a minimum value of *e* varies directly as the value of *u*.

XVI

Conclusion

It is evident that we have left unconsidered many important problems in the preceding pages. Our only excuse is that we have made no pretensions to completeness. There is, however, one point which may be touched upon before closing.

There is a concept, developed in a field of psychological endeavor widely separated from the one we have been considering, which is nevertheless applicable to our present problems. We refer to the well-known *interval of uncertainty* in psychophysics. Every stimulus to which an organism reacts has, for that organism, an interval of uncertainty, or, better still, an *interval of equivalence.* That is to say, the stimulus or situation may be changed by a certain amount either in one direction or another as regards its form or intensity and still evoke from the organism the same response under otherwise identical conditions. Klüver (21) in his recent research has forcefully expressed the importance of this point and has used for it the term "equivalence of stimuli."

This is a point which hitherto has not been kept in mind sufficiently either in animal or human research. It is, nevertheless, a fact of pivotal importance (10) and one which, when clearly kept in mind, solves at one stroke a number of questions which have arisen in consequence of a too atomistic view of the neural processes which are involved in behavior. Thus it is that many investigators have wondered why an animal behaves adequately in a situation which, microscopically considered, he has never faced previously. Clearly, we are touching here upon the very old but ubiquitous problem of similarity and generalization or abstraction. (See reference 10.)

In conclusion, we are of the opinion that it would miss the mark to consider the preceding discussions as an elaboration of a *conditioned response* theory of learning. The conditioned

response itself is only a special case of modifiability in general. Thus, in the present theory, it would have no meaning to say that the conditioned reflex is the unit of learning process. The learning process, in our opinion, *has no structural units.*

We believe, therefore, that we have not given a " mechanical " theory of learning in the preceding pages. Contrariwise, we trust to have made sufficiently clear that the neural processes determining behavior must be essentially dynamic, in that any given " component " in a complex temporal or simultaneous stimulation leaves in the cerebral mass " traces " which are determined by the process as a whole in the true " Gestaltean " sense. We have attempted, both implicitly and explicitly, to point out the futility of attempts to envisage neural processes as being constituted of isolated paths of the telephone type which conduct impulses independently of past and contemporary excitations, and whose after-effects cannot be influenced by succeeding conditions.

There are a number of interesting deductions which are to be made from the foregoing considerations and which are amenable to experimental test. Of these, we have mentioned but a few. It is hoped that, if the discussion has any plausibility, many more of these will become apparent to the reader.

REFERENCES

1. Adrian, E. D. The basis of sensation. New York: Norton, 1928, Pp. 122.
2. Blair, H. A. On intensity-time relations for stimulation by electric currents. *J. Gen. Physiol.*, 1932, 15, 709–729.
3. Blodgett, H. C. The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, 4, 113–134.
4. Buel, J. (Article to appear in *Univ. Calif. Publ. Psychol.*)
5. De Camp, J. E. Relative distance as a factor in the white rat's selection of a path. *Psychobiol.*, 1920, 2, 245–253.
6. Elliott, M. H. The effect of change of " drive " on the maze performance. *Univ. Calif. Publ. Psychol.*, 1929, 4, 185–188.
7. Erlanger, J., Bishop, G. H., and Gasser, H. S. Distortion of action potentials as recorded from the nerve surface. *Amer. J. Physiol.*, 1926, 78, 592–609.
8. Forbes, A. The interpretation of spinal reflexes in terms of present knowledge of nerve conduction. *Physiol. Rev.*, 1922, 2, 361–414.
9. Gasser, H. S., and Graham, H. T. The end of the spike potential of nerve and its relation to the beginning of the after-potential. *Amer. J. Physiol.*, 1932, 101, 316–330.
10. Gengerelli, J. A. Studies in abstraction with the white rat. *J. Genet. Psychol.*, 1930, 38, 171–201.

11. GENGERELLI, J. A. The principle of maxima and minima in animal learning. *J. Comp. Psychol.*, 1930, 11, 193–236.
12. GENGERELLI, J. A. Similarity and retroaction. (To appear in *J. Exp. Psychol.*)
13. GERARD, R. W. Nerve conduction in relation to nerve structure. *Quarterly Rev. Biol.*, 1931, 6, 59–83.
14. GUTHRIE, E. R. Conditioning as a condition of learning. *Psy. Rev.*, 1930, 37, 412–428.
15. HECHT, S. The nature of the photoreceptor process. In The Foundations of Experimental Psychology, ed. by Carl Murchison, Worcester, 1929. Pp. 216.
16. HILGARD, E. R. Conditioned eyelid reactions to a light stimulus based on the reflex wink to sound. *Psychol. Monog.*, 1931, 41, No. 184. Pp. 50.
17. HOAGLAND, H. Specific afferent impulses and cutaneous sensitivity. *J. Gen. Psychol.*, 1932, 6, 276–295.
18. HULL, C. L. The goal-gradient hypothesis and maze learning. *Psychol. Rev.*, 1932, 39, 25–43.
19. HUNTER, W. S. The behavior of raccoons in a double alternation temporal maze. *J. Genet. Psychol.*, 1928, 35, 374–389.
20. KLEITMAN, N., and CRISLER, G. A quantitative study of the salivary conditioned reflex. *Amer. J. Physiol.*, 1926, 79, 571–614.
21. KLÜVER, H. The equivalence of stimuli in the behavior of monkeys. *J. Genet. Psychol.*, 1931, 39, 3–27.
22. KUO, Z. Y. The nature of unsuccessful acts and their order of elimination. *J. Comp. Psychol.*, 1922, 2, 1–27.
23. LILLIE, R. S. Protoplasmic action and nervous action. Chicago, 1932.
24. PAVLOV, I. P. Conditioned reflexes. London, 1927.
25. PAVLOV, I. P. The reply of a physiologist to psychologists. *Psychol. Rev.*, 1932, 39, 91–127.
26. ROBINSON, E. S. The 'similarity' factor in retroaction. *Amer. J. Psychol.*, 1927, 39, 297–312.
27. SWITZER, ST. C. A. Backward conditioning of the lid reflex. *J. Exper. Psychol.*, 1930, 13, 76–97.
28. TOLMAN, E. C. Sign-Gestalt or conditioned reflex. *Psychol. Rev.*, 1933, 40, 246–255.
29. TOLMAN, E. C., and HONZIK, C. H. Degrees of hunger, reward and non-reward, and maze learning in rats. *Univ. Calif. Publ. Psychol.*, 1930, 4, 241–257.
30. TRAVIS, L. E., and HERREN, R. Y. The relation of electrical changes in the brain to reflex activity. *J. Comp. Psychol.*, 1931, 12, 23–39.
31. TROLAND, L. T. The principles of psychophysiology, II. New York: Van Nostrand, 1930. Pp. XXI+397.
32. YOSHIOKA, J. Weber's law in the discrimination of maze distance by the white rat. *Univ. Calif. Publ. Psychol.*, 1929, 4, 155–184.